



## REQUIEM-2D Methodology

### Spatially Resolved Stellar Populations of Massive Lensed Quiescent Galaxies from Hubble Space Telescope 2D Grism Spectroscopy

Akhshik, Mohammad; Whitaker, Katherine E.; Brammer, Gabriel; Mahler, Guillaume; Sharon, Keren; Leja, Joel; Bayliss, Matthew B.; Bezanson, Rachel; Gladders, Michael D.; Man, Allison; Nelson, Erica J.; Rigby, Jane R.; Rizzo, Francesca; Toft, Sune; Wellons, Sarah; Williams, Christina C.

*Published in:*  
Astrophysical Journal

*DOI:*  
[10.3847/1538-4357/abac62](https://doi.org/10.3847/1538-4357/abac62)

*Publication date:*  
2020

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)

*Citation for published version (APA):*  
Akhshik, M., Whitaker, K. E., Brammer, G., Mahler, G., Sharon, K., Leja, J., Bayliss, M. B., Bezanson, R., Gladders, M. D., Man, A., Nelson, E. J., Rigby, J. R., Rizzo, F., Toft, S., Wellons, S., & Williams, C. C. (2020). REQUIEM-2D Methodology: Spatially Resolved Stellar Populations of Massive Lensed Quiescent Galaxies from Hubble Space Telescope 2D Grism Spectroscopy. *Astrophysical Journal*, 900(2), [184].  
<https://doi.org/10.3847/1538-4357/abac62>



# REQUIEM-2D Methodology: Spatially Resolved Stellar Populations of Massive Lensed Quiescent Galaxies from Hubble Space Telescope 2D Grism Spectroscopy

Mohammad Akhshik<sup>1</sup>, Katherine E. Whitaker<sup>2,3</sup>, Gabriel Brammer<sup>3,4</sup>, Guillaume Mahler<sup>5</sup>, Keren Sharon<sup>5</sup>, Joel Leja<sup>6,16</sup>, Matthew B. Bayliss<sup>7</sup>, Rachel Bezanson<sup>8</sup>, Michael D. Gladders<sup>9,10</sup>, Allison Man<sup>11</sup>, Erica J. Nelson<sup>6,17</sup>, Jane R. Rigby<sup>12</sup>,  
 Francesca Rizzo<sup>13</sup>, Sune Toft<sup>3,4</sup>, Sarah Wellons<sup>14</sup>, and Christina C. Williams<sup>15,16</sup>

<sup>1</sup> Department of Physics, University of Connecticut, Storrs, CT 06269, USA; [mohammad.akhshik@uconn.edu](mailto:mohammad.akhshik@uconn.edu)

<sup>2</sup> Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA

<sup>3</sup> Cosmic Dawn Center (DAWN), Denmark

<sup>4</sup> Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen, Denmark

<sup>5</sup> Department of Astronomy, University of Michigan, 1085 South University Ave, Ann Arbor, MI 48109, USA

<sup>6</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden St. Cambridge, MA 02138, USA

<sup>7</sup> Department of Physics, University of Cincinnati, Cincinnati, OH 45221, USA

<sup>8</sup> Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA

<sup>9</sup> Department of Astronomy & Astrophysics, The University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA

<sup>10</sup> Kavli Institute for Cosmological Physics at the University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA

<sup>11</sup> Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St George Street, Toronto ON, M5S 3H4, Canada

<sup>12</sup> Observational Cosmology Lab, NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD 20771, USA

<sup>13</sup> Max-Planck Institute for Astrophysics, Karl-Schwarzschild Str 1, D-85748 Garching, Germany

<sup>14</sup> Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA

<sup>15</sup> Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

Received 2019 July 5; revised 2020 August 3; accepted 2020 August 3; published 2020 September 15

## Abstract

We present a novel Bayesian methodology to jointly model photometry and deep Hubble Space Telescope 2D grism spectroscopy of high-redshift galaxies. Our `requiem2d` code measures both unresolved and resolved stellar populations, ages, and star formation histories (SFHs) for the ongoing RESolving QUIEscent Magnified (REQUIEM) Galaxies Survey, which targets strong gravitationally lensed quiescent galaxies at  $z \sim 2$ . We test the accuracy of `requiem2d` using a simulated sample of massive galaxies at  $z \sim 2$  from the Illustris cosmological simulation and find that we recover the general trends in SFH and median stellar ages. We further present a pilot study for the REQUIEM Galaxies Survey: MRG-S0851, a quintuply imaged, massive ( $\log M_*/M_\odot = 11.02 \pm 0.04$ ) red galaxy at  $z = 1.883 \pm 0.001$ . With an estimated gravitational magnification of  $\mu = 5.7^{+0.4}_{-0.2}$ , we sample the stellar populations on 0.6 kpc physical size bins. The global mass-weighted median age is constrained to be  $1.8^{+0.3}_{-0.2}$  Gyr, and our spatially resolved analysis reveals that MRG-S0851 has a flat age gradient in the inner 3 kpc core after taking into account the subtle effects of dust and metallicity on age measurements, favoring an early formation scenario. The analysis for the full REQUIEM-2D sample will be presented in a forthcoming paper with a beta release of the `requiem2d` code.

*Unified Astronomy Thesaurus concepts:* Galaxy formation (595); High-redshift galaxies (734); Galaxy stellar content (621); Galaxy quenching (2040); Strong gravitational lensing (1643)

## 1. Introduction

Our understanding of galaxies a few billion years after the Big Bang has dramatically improved over the last few decades. It is now well-established that galaxies follow a bimodal color distribution in both the low- and high-redshift universe (e.g., Strateva et al. 2001; Whitaker et al. 2011), including a population of old, red, more massive quiescent galaxies and a population of young, blue, less massive star-forming galaxies. Star-forming and quiescent galaxies can be identified by their location in the star formation rate (SFR) versus stellar mass plane, where star-forming populations form a sequence with a relatively low scatter (e.g., Speagle et al. 2014; Whitaker et al. 2014a) and quiescent populations lie well below the average relation. The number density of massive quiescent galaxies rapidly increased at early times, comprising up to half of the total massive galaxy population by  $z \sim 2$  (Kriek et al. 2006; Brammer et al. 2011; Muzzin et al. 2013). Moreover,

observations show these quiescent galaxies to be remarkably compact relative to star-forming galaxies with similar stellar masses at a given redshift (e.g., van Dokkum et al. 2008; van der Wel et al. 2014), with only the most massive galaxies ( $\log M_*/M_\odot > 11.3$ ) having similar size distributions as the star-forming population (Mowla et al. 2018).

Despite the tremendous progress in understanding the population of  $z \sim 2$  massive galaxies, usually presented in empirical correlations like the SFR–stellar mass correlation of star-forming galaxies described above, the physical mechanisms responsible for quenching star-forming galaxies remain unknown. Spatially resolved spectroscopy and imaging hold the power to address these fundamental questions. Simulations suggest that stellar age and specific star formation rate (sSFR) gradients can constrain the theoretical formation scenarios for high-redshift massive quiescent galaxies (e.g., Tacchella et al. 2015, 2016; Wellons et al. 2015). However, the low spatial resolution of near- and mid-infrared imaging, along with the high stellar density of quiescent galaxies, mostly limit the studies to the spatially unresolved data with relatively less constraining power to distinguish between theoretical models

<sup>16</sup> NSF Astronomy and Astrophysics Postdoctoral Fellow.

<sup>17</sup> Hubble Fellow.

(e.g., Williams et al. 2017; Abramson et al. 2018; Belli et al. 2019; Estrada-Carpenter et al. 2020). Strong gravitational lensing offers a solution for this challenge, as it magnifies distant galaxies and boosts their signal-to-noise ratio (S/N). Furthermore, the actual unlensed morphology can be reconstructed accurately with a detailed lensing model (e.g., Sharon et al. 2012, 2020).

Strong gravitationally lensed galaxies are discovered and studied extensively in the literature (e.g., Williams & Lewis 1996; Yee et al. 1996; Allam et al. 2007; Smail et al. 2007; Siana et al. 2008; Belokurov et al. 2009; Lin et al. 2009; Koester et al. 2010; Sharon et al. 2012; Gladders et al. 2013), with many cases of spatially resolved stellar population analyses in star-forming galaxies (e.g., Stark et al. 2008; Swinbank et al. 2009; Jones et al. 2010; Leethochawalit et al. 2016). Despite their rarity, a number of ground-based spectroscopic studies (e.g., Keck/MOSFIRE, Magellan/FIRE, and VLT/X-Shooter) of massive quiescent galaxies have steadily accumulated within the literature (Muzzin et al. 2012; Geier et al. 2013; Newman et al. 2015, 2018a, 2018b; Hill et al. 2016; Toft et al. 2017; Ebeling et al. 2018). However, ground-based spatial resolution is insufficient to resolve spectroscopic signatures of the stellar populations of all but perhaps the most strongly lensed objects (Newman et al. 2015).

The high spatial and low spectral resolution of grism spectroscopy with the Hubble Space Telescope (HST)/Wide Field Camera 3 (WFC3) enables measuring both the unresolved and/or resolved stellar populations (e.g., van Dokkum & Brammer 2010; Brammer et al. 2012a; Whitaker et al. 2013, 2014b; Estrada-Carpenter et al. 2019, 2020; D’Eugenio et al. 2020). In particular, Abramson et al. (2018) use WFC3/G141 grism spectroscopy and multiwavelength HST imaging to study the spatially resolved stellar populations of four massive galaxies at  $z \sim 1.3$ , finding a link between bulge mass function and the shape of the star formation history. Similar comprehensive studies of massive quiescent galaxies at higher redshifts demonstrate that it is feasible to reconstruct star formation histories (SFHs) based on a joint spectrophotometric HST analyses (Morishita et al. 2018, 2019). While the measured metallicities of quiescent galaxies at  $z \sim 2$  are generally found to be similar to local early-type galaxies (Morishita et al. 2019), there exist a few lensed quiescent galaxies with lower metallicities that suggest a mechanism other than dry minor mergers would be necessary to explain their chemical enrichment (Morishita et al. 2018).

In this paper, we present our methodology developed to jointly fit HST and Spitzer-IRAC spectrophotometric data in preparation for the analysis of the full REsolving QUIEscent Magnified (REQUIEM) galaxy survey (HST-GO-15633). While it is possible to constrain stellar population properties by analyzing spatially resolved<sup>18</sup> spectroscopic and photometric data separately, we perform a joint spectrophotometric fit because, using a joint fit, we can optimally use all spectrophotometric data (e.g., Newman et al. 2014) and infer

all parameters within a single framework. As a single set of assumptions is applied in this joint fitting, it is also easier to understand and address potential biases and systematics.

We briefly introduce the REQUIEM galaxy survey in Section 2, and illustrate our method using a pilot target, MRG-S0851 (Sharon et al. 2020). The HST and Spitzer data reductions are presented in Section 3. The methodology to jointly fit photometry and spectroscopy is presented in Section 4. We discuss inferring ages and star formation histories in Section 5, testing the inferred parameters using a sample of massive quiescent and star-forming galaxies selected from the Illustris simulation. In Section 6, we present the first results from REQUIEM-2D grism spectroscopy for our pilot target, MRG-S0851. In Appendices A and B, we discuss the details of the lensing model and the morphological measurements of MRG-S0851.

In this paper, we adopt a standard simplified  $\Lambda$ CDM cosmology with  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . We assume the Chabrier (2003) initial mass function. All magnitudes are reported in the AB system.

## 2. REQUIEM-2D Galaxy Survey

Capitalizing on the decade-long hunt for strong lensed quiescent galaxies at  $z > 1.5$  and the slitless spectroscopic capabilities of HST, the REQUIEM-2D galaxy survey targets eight strongly lensed quiescent galaxies spanning redshifts of  $1.6 < z < 2.9$ , stellar masses of  $10.4 < \log M_*/M_\odot < 11.7$ , and sSFRs of  $\log \text{sSFR}_{100 \text{ Myr}}/[\text{yr}^{-1}] < -10.3$  (HST-GO-15633).

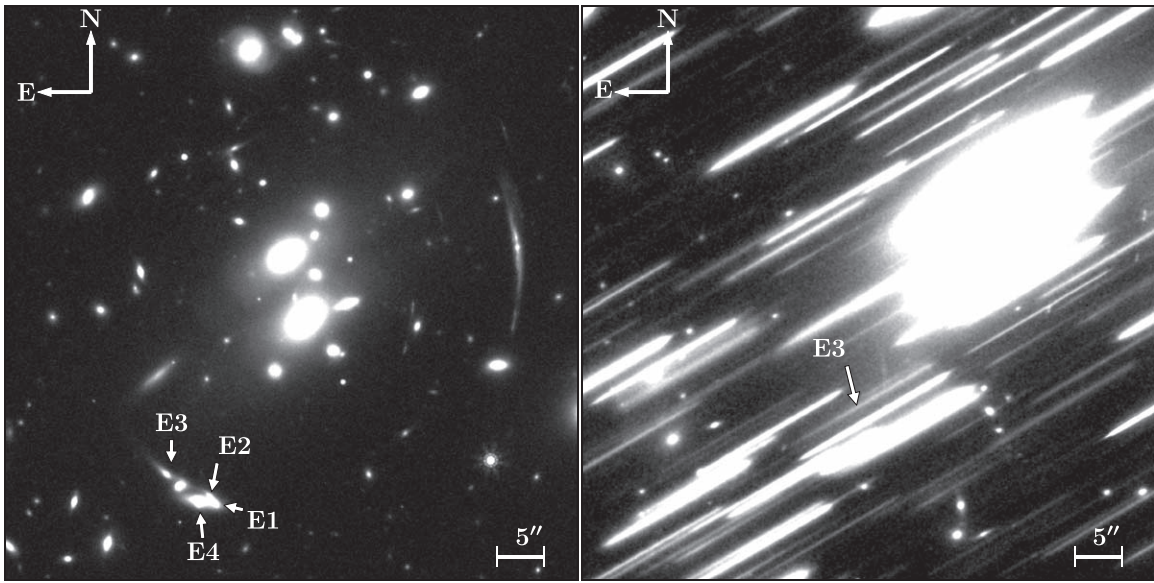
Next, we briefly introduce the targets comprising the REQUIEM-2D survey, with the pilot target MRG-S0851 described in further detail in Section 6. Our sample includes the following.

1. MRG-M1341: a highly magnified  $\mu \sim 30$  galaxy at  $z = 1.6$  (Ebeling et al. 2018) (15 orbits of WFC3/G141).
2. MRG-S0851: a massive lensed red galaxy at  $z = 1.88$ , with centrally concentrated rest-frame UV flux (12 orbits of WFC3/G141; presented in this paper).
3. MRG-M0138: a massive and bright target at  $z = 1.95$  with  $\log M/M_\odot = 11.7$  and  $H_{F160W} = 17.3$  (Newman et al. 2018a) (six orbits of WFC3/G141).
4. MRG-P0918 and MRG-S1522: relatively young quiescent galaxies at  $z = 2.36$  and  $z = 2.45$ , respectively, with ages of 0.5–0.6 Gyr (Newman et al. 2018a) (seven orbits of WFC3/G141 each).
5. MRG-M2129: a rotationally supported quenched galaxy at  $z = 2.1$  (Toft et al. 2017) (five orbits of WFC3/G141).
6. MRG-M0150: a dispersion-dominated ( $V/\sigma = 0.7 \pm 0.2$ ) massive quiescent galaxy at  $z = 2.6$  (Newman et al. 2015) (five orbits of WFC3/G141).
7. MRG-S0454: the most compact  $r_{\text{eff}} \sim 0.3 \text{ kpc}$  target of the REQUIEM-2D survey with the highest redshift of  $z = 2.9$  (A. Man et al. 2020, in preparation) (12 orbits of WFC3/G141).

The number of HST bands available for the REQUIEM targets ranges from a minimum of five filters to a maximum of 16. All targets have photometric coverage from  $\sim 1000 \text{ \AA}$  to  $\sim 15000 \text{ \AA}$  in rest-frame wavelength, and grism G141 coverage varies from rest-frame wavelengths of  $\sim 2900\text{--}4200 \text{ \AA}$  for the target with the highest redshift to  $\sim 4400\text{--}6300 \text{ \AA}$  for the target with the lowest redshift. The imaging data used herein for the

<sup>18</sup> We caution that the term spatially “resolved” in nearby galaxies is reserved for an observational study that can resolve stars down to at least  $\lesssim \mathcal{O}(10^6)$  stars per pixel (e.g., Cook et al. 2019). This limit corresponds to distances of  $< 1 \text{ Mpc}$  with HST detectors—and nominally, even individual stellar clusters can be identified in “spatially resolved” studies for nearby targets and surveys (e.g., Johnson et al. 2012, 2015). Our targets are well beyond this limit, but we can still resolve stellar populations down to a fraction of kpc scale, and we therefore use the term spatially “resolved” to refer to our study, noting the conceptual difference in the terminology used for nearby and  $z \sim 2$  galaxies.





**Figure 1.** HST imaging and spectroscopy of MRG-S0851. On the left, the drizzled mosaics of WFC3  $H_{F160W}$  filter are shown. The four main images of MRG-S0851, a pilot target from REQUIEM survey, are indicated by white arrows. The fifth image, E5, is right next to the subcluster lens, but it is not shown here. We indicate E5 in Figure A2, where the light profile of subcluster lens is modeled and subtracted from the image. On the right, the drizzled mosaic for  $\sim 6$  orbits of the WFC3/G141 data is shown. White arrow indicates the grism spectrum of E3. E3 is the cleanest and the brightest image of the system, MRG-S0851 in the H-band, with our HST grism observations optimized to reduce contamination for E3 at the expense of losing E1, E2, and E4.

test target, MRG-S0851, consists of five HST bands and two Spitzer bands (see Section 6).

### 3. Data Reduction and Analysis

#### 3.1. HST Grism Spectroscopy

The REQUIEM-2D HST observations are designed following the 3D-HST standard (Brammer et al. 2012b), including a shorter exposure with a WFC3/IR imaging filter immediately before/after two longer WFC3/IR G141 exposures. The “Grism redshift and line analysis software for space-based slitless spectroscopy,” or *Grizli*, is used for the data reduction analysis (Brammer 2016). *Grizli* is specifically designed for manipulating HST slitless spectroscopic observations, and it serves for the data reduction herein.

Astrometric calibrations of the WFC3-IR and WFC3-UVIS images are performed in two steps within *Grizli*. In the first step, the relative astrometry is set by aligning all available exposures in each filter together. The Pan-STARRS catalog (Flewelling et al. 2016) is then used to density match the detected objects. The absolute astrometric registration is finally improved by adopting the Gaia-DR2 catalog (Gaia Collaboration et al. 2018; Lindegren et al. 2018).

*Grizli* matches the world coordinate system (WCS) of the grism exposures with already registered WFC3/IR exposures and subtracts the sky background after reducing and calibrating the grism exposures. All of the exposures are drizzled together using the *AstroDrizzle* package (Avila & Hack 2012). Figure 1 shows the final product for MRG-S0851, a pilot target from the REQUIEM galaxy survey (see Section 6).

WFC3/G141 grism produces dispersed spectra of every object within the field of view of the instrument. Without slits, however, spectra of nearby objects overlap. To analyze the 2D grism spectrum of an object of interest, contamination by other objects must be removed. Here, we adopt an iterative algorithm within *Grizli* to remove contamination. First, 2D grism models are generated for all objects, assuming a flat spectrum

that we refine iteratively. In subsequent steps, we concentrate on the region surrounding the primary science target, which extends roughly a factor of five times beyond the largest spatial extent of the main science target. The grism model of all objects is refined in this surrounding region by using second- and fifth-degree polynomials as spectral templates. A linear combination of a flexible set of spectral templates, built in *Grizli* to constrain redshifts (Brammer et al. 2008), is finally fit to improve the quality of the model. Figure 2 demonstrates the result of this procedure for the pilot target, MRG-S0851.

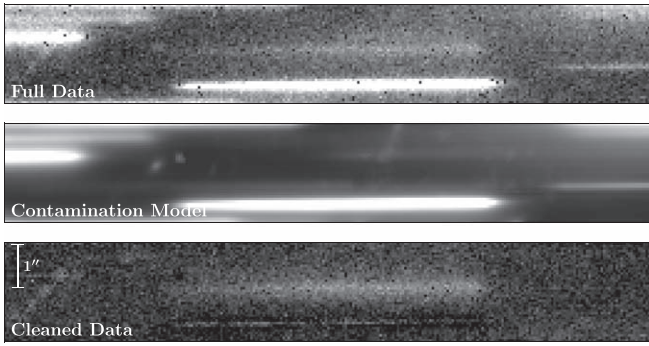
#### 3.2. Photometric Measurements

To perform the joint spectrophotometric analysis, we first construct a photometric catalog, largely following Whitaker et al. (2011) and Skelton et al. (2014). We refer the reader to these papers for a more in-depth discussion on the methodology adopted.

##### 3.2.1. HST Photometry

To detect sources, we first construct a noise-equalized image by multiplying the  $H_{F160W}$  mosaic with the square root of the corresponding weight map. We then run *Source Extractor* (Bertin & Arnouts 1996) on this image. The detection threshold is set at  $1.8\sigma$ , the deblending threshold at 32, with a minimum contrast of 0.0001 and a minimum area of 14 pixels.

To create the point-spread functions (PSF), a stellar sequence is identified by considering the ratio of a small aperture ( $0''.5$ ) flux to a large aperture ( $2''$ ) flux for each band. Stars form a tight sequence close to unity, making them easily identifiable above a certain threshold in magnitude. A  $5''$  postage stamp cutout of each bright star is created. An average PSF is calculated after centering and normalizing the stamps. The PSF matching is performed using a kernel that convolves each PSF to match the  $H_{F160W}$  PSF as a reference, since it has the largest full width at half maximum (FWHM) of  $0''.18$ . To obtain the kernel, we use custom codes that fit a set of Hermite



**Figure 2.** A  $15''.6 \times 2''.4$  cutout region of a single G141 exposure centered at E3, the brightest image of the pilot target MRG-S0851. In the top panel, we show the full data. Second panel shows the contamination model, which we obtain by iteratively fitting polynomial spectral templates to the grism spectra of all objects. Third panel shows cleaned grism data of E3.

polynomials weighted by Gaussian two-dimensional profiles to the Fourier transform of the stacked stars. The PSF homogenization is accurate within a percent level.

Next, `Source Extractor` is run in the dual-image mode with the noise-equalized image of  $H_{F160W}$  as a detection image and the PSF-matched mosaic of interest as a measurement image, including the weight maps of the PSF-matched mosaics as well. The photometry is calculated adopting an aperture of  $1''.5$  diameter for all but the most extended strong lensed sources. This is about a factor of two larger than the aperture size adopted in earlier works, but justified when analyzing strong gravitationally lensed sources with linear magnifications of  $\mu \sim 3\text{--}6$  (e.g.,  $0''.7 \sim 1''.5/\sqrt{\mu}$ ), since the larger aperture in the image plane of lensed target effectively covers the same physical region in the source plane as the smaller aperture would cover for unlensed targets.

The curve of growth of the  $H_{F160W}$  filter is used to correct the AUTO flux value reported by `Source Extractor` for the amount of light falling outside the Kron radius (Kron 1980). This correction factor is the ratio of the total flux of a point source in  $H_{F160W}$  to the flux enclosed in the Kron radius (e.g., Skelton et al. 2014).

Realistic uncertainties are estimated by placing apertures in empty regions across the image and calculating the noise properties directly from the images in lieu of using the standard weight maps, noting that the drizzling process correlates the pixels, and as a result the uncertainty inferred from the weight maps is underestimated (e.g., Casertano et al. 2000). More details can be found in Section 3.5 of Whitaker et al. (2011) and Section 3.4 of Skelton et al. (2014).

### 3.2.2. Spitzer/IRAC Photometry

To obtain photometric measurements from the low-resolution Spitzer observations, we use the Multi-resolution Object PHotometry ON Galaxy Observations code (MOPHONGO; Labbé et al. 2006; Wuyts et al. 2007). MOPHONGO makes two-dimensional models for different objects in the field and uses them to deblend and measure fluxes, taking into account the difference in PSF between Spitzer and HST images.

Following Whitaker et al. (2011), Spitzer photometric fluxes are measured using  $3''$  diameter apertures size, applying photometric corrections using the  $H_{F160W}$  curve of growth. While they have poor resolution, the photometric measurements of the two Spitzer IRAC channels play a crucial role in

the modeling of stellar populations, owing to the extended wavelength coverage into the rest-frame near-infrared at  $z \sim 2$  that helps to constrain the dust (e.g., Muzzin et al. 2008.)

## 4. Methodology to Fit the Age and SFH of the Stellar Populations

In this section, we discuss the methodology used by the `requiem2d` software package to combine all spectroscopic and photometric data and constrain the age and SFHs of unresolved and resolved stellar populations. An overview of the main aspects of our methodology is presented in Section 4.1. We then outline our approach to model dust and metallicity in Section 4.2, before formally introducing the elements of the full model in Section 4.3. A discussion of priors and the computational Bayesian approach can be found in Section 4.4.

### 4.1. Overview of Methodology

The `requiem2d` package adopts a nonparametric framework to model SFHs, avoiding any assumptions about their functional form (see Section 4.4 for a discussion of SFH priors). Joint spectrophotometric fitting is particularly important for a robust analysis of the stellar populations, with the longer wavelength baseline of the photometry helping to constrain dust and the higher spectral resolution grism spectroscopy providing more robust constraints on redshift, age, and metallicity by constraining spectral absorption lines.

We adopt a nonparametric approach to analyze SFHs, specifically modeling the composite stellar population (CSP) of the targets as a linear combination of simple stellar populations (SSPs) (e.g., Heavens et al. 2004; Ocvirk et al. 2006; Panter et al. 2007; Tojeiro et al. 2007; Kelson et al. 2014; Leja et al. 2017; Dressler et al. 2018; Morishita et al. 2019), which is used to constrain the “weights” of each SSP, denoted herein by  $x$ . The secondary parameters such as age and SFR are then calculated using these weights. This methodology, in principle, is similar to the approach adopted in EAZY (Brammer et al. 2008), where one fits a linear combination of templates to photometric data to constrain the redshift. Here, we fit a linear combination of the SSP templates with varying ages to the low-resolution spectroscopic and photometric data.

To generate SSPs, we use the dust and metallicity posteriors obtained by fitting the photometric data alone (Section 4.2). We then refit the full spectrophotometric data to infer ages and SFHs using these SSPs (Section 4.3). In the remainder of this Section, we discuss data preparation steps (Sections 4.1.1 and 4.1.2).

#### 4.1.1. Defining the Spatial Bins

To study the spatially resolved stellar populations for lensed targets, we define spatial bins for each grism exposure separately, using the corresponding direct WFC3/IR image with the same pixel scale and orientation as the grism exposure. “Rows of pixels” are defined parallel to the dispersion angle  $P_\theta$ . We identify the row that has the pixel with the highest flux in the image and add two adjacent pixel rows to define the central bin.

On either side of the center bin, two bins are defined that are 3–4 pixel rows wide, respectively. Depending on the magnification of the main science target, either new subsequent spatial bins are added or the rest of the pixels on each side are

grouped to define the final outer bins. These other bins include the pixels corresponding to the low-S/N portion of the extended light profiles.

With a pixel size of  $0''.06$ , the central bins range from  $0''.18$  to  $0''.24$  wide. Lens models are used to determine the source-plane position of the defined spatial bins. For our pilot study of MRG-S0815 (Section 6), we probe the age gradient in the inner radius of  $\sim 1.8$  kpc at an average spatial resolution of  $\sim 0.6$  kpc.

#### 4.1.2. Preparing the Data

Grism spectra are analyzed in the native 2D space, limiting to grism pixels with a minimum S/N of 3. We also only include the grism pixels with less than 10% contamination by adjacent objects.

It is not trivial to spatially resolve the Spitzer/IRAC bands. In our final joint fitting, we therefore adopt a conservative approach by requiring that the total IRAC fluxes of all spatial bins match the global measured IRAC flux of the object. We discuss other complications of not having resolved rest-frame near-infrared (IRAC) fluxes in Section 4.2.2.

The grism spectra of the bins overlap in 2D space, making it impossible to extract the 2D grism spectrum for each bin individually unless we have the best model for the other bins. We therefore construct a model for each bin individually with *Grizli*, then add all of the bins' models together to get a model for the whole galaxy.

We could, in principle, model all grism exposures individually and compare them with the observed grism exposure. However, to reduce the computational cost, we use drizzled grism images in our analyses, constructed by combining all grism exposures of each dispersion angle.

### 4.2. Measuring Dust and Metallicity

The main goal when using *requiem2d* is to constrain the ages and star formation histories of massive quiescent galaxies at  $z \sim 2$ , treating dust and metallicity as nuisance parameters. While both of these parameters are degenerate with age, they are not well-constrained by the relatively short wavelength coverage and low spectral resolution of grism spectroscopy. Hence, our strategy is to analyze the problem in two steps:

1. Photometric data are fitted alone using *Prospector- $\alpha$*  (Leja et al. 2017) to obtain the posterior of dust, metallicity, and other relevant parameters of stellar populations, such as the stellar mass (Section 4.2.1). The posteriors of dust and metallicity are subsequently used to generate SSPs (Section 4.2.2).
2. Joint fits of photometric and spectroscopic data are performed using *requiem2d* code to constrain the age and SFHs of the stellar populations, using the SSPs generated at the first step (Section 4.3).

#### 4.2.1. *Prospector- $\alpha$* Fit to Resolved and Global Photometric Data

Following the same steps and assumptions of Leja et al. (2019b), the photometric data are fit using *Prospector- $\alpha$* . In particular, we adopt a nonparametric approach to model SFHs, imposing the continuity prior (Leja et al. 2019a, 2019b). The continuity prior disfavors unphysical jumps in SFH, i.e., episodes of rejuvenation and quenching, and it leads to a more physical and smoother SFH (Leja et al. 2019a). We refer the

reader to Leja et al. (2019a) for further discussion of different priors of SFH, noting that the prior that we assume in our joint fitting is similar to the continuity prior (see Equation (3)).

*Prospector- $\alpha$*  adopts the Kriek & Conroy (2013) dust model, which is based on the parameterization of dust attenuation by Noll et al. (2009). In this model, the strength of the 2175Å UV bump is correlated with the dust slope. Therefore, the free parameters are *dust\_index* (dust slope) as well as two dust attenuation parameters, *dust1* and *dust2* for the stellar populations younger and older than  $10^7$  yr, respectively. We note that the *dust\_index* parameter controls the slope of dust attenuation curve, and for positive values, the attenuation curve will be flatter than the Calzetti law (e.g., Kriek & Conroy 2013, Figure 1), leading to comparatively less UV attenuation and more near-IR attenuation. For the negative values of *dust\_index*, the opposite holds.

We use the Mesa Isochrones and Stellar Tracks (MIST; Choi et al. 2016) to generate the SSPs using Flexible Stellar Populations Synthesis models (FSPS; Conroy et al. 2009; Conroy & Gunn 2010). The stellar metallicity is therefore measured relative to the solar abundance, as defined in Table 1 of Choi et al. (2016), and it is constrained by *Prospector- $\alpha$*  based on the UV to optical to near-IR ratios of the SED (see Figure 3 of Leja et al. 2017).

#### 4.2.2. Priors to Generate SSPs for Spatially Resolved and Global Joint Fit

We include all global photometric measurements (HST and Spitzer) in the *Prospector- $\alpha$*  fit and use the resulting posterior as a prior in the joint fitting.

For fitting the spatially resolved stellar populations, we calculate the spatially resolved photometric fluxes for the HST bands by summing the flux for all pixels in each bin. To estimate the photometric uncertainty and to be sure that the correlated pixel noises are accounted for, we follow Whitaker et al. (2011) and Skelton et al. (2014), where the uncertainty is scaled by a power-law function of aperture sizes approximated by  $\sqrt{N}$ , where  $N$  is the number of pixels in each bin. We fit the resolved HST photometry and error of each spatial bin using *Prospector- $\alpha$* .

To take into account the dust and metallicity uncertainties in the spatially resolved joint fit, we have two choices of priors to generate SSPs, corresponding to two *Prospector- $\alpha$*  fits. The first prior is defined using the spatially resolved *Prospector- $\alpha$*  fit, while the second prior is defined by the global *Prospector- $\alpha$*  fit for all spatial bins (see Section 4.3.1 for the details regarding inclusion of the dust and metallicity uncertainties for generating SSPs). The first prior is tuned to the resolved HST bands of the spatial bins, but the corresponding *Prospector- $\alpha$*  fit does not include IRAC channels. The second prior is not tuned to the individual bins, but it does include the IRAC channels 1 and 2. As there is no clear preference *a priori*, we perform our spatially resolved joint fitting adopting both of these priors. All of the fits being performed, with their SSP priors and included data, are summarized in Table 1.

### 4.3. Elements of the *requiem2d* Full Model

In this section, we discuss the building blocks of our Bayesian model: the elements of the regression model, the prior, and the likelihood distributions.



**Table 1**  
Five Different Fits Performed in Our Analyses, Indicating Software Used, Including Data and SSP Prior if Used

Fit	Software	SSP Prior	Grism	Res HST Phot	Unres HST Phot	Spitzer Phot
Global Phot	Prospector- $\alpha$	×	×	×	✓	✓
Resolved Phot	Prospector- $\alpha$	×	×	✓	×	×
Global Spec+Phot	requiem2d	Global Phot	✓	×	✓	✓
Resolved Spec+Phot	requiem2d	Global Phot	✓	✓	×	✓
Resolved Spec+Phot	requiem2d	Resolved Phot	✓	✓	×	✓

**Note.** Phot is a shorthand for photometry, Spec stands for spectroscopy, and Res and Unres stand for Resolved and Unresolved, respectively.

**Table 2**  
Elements of the requiem2d Generalized Regression Model Used in the Joint SpectroPhotometric Fit

Weights	Predictors	Description
$[x_{ij}]_{M \times N}$	$[A_{s,ijl}]_{M \times N \times X}$ $[A_{p,ijr}]_{M \times N \times P}$	$A_s$ and $A_p$ are SSP templates $M$ is the total number of spatial bins, $N$ is the total number SSPs for each age, $X$ is the total number of G141 pixels.
$P$ is the total number of photometric bandpasses. $[x_{em,iq}]_{M \times 4}$	$[A_{em,iql}]_{M \times 4 \times X}$	The weight $x_{em}$ of the emission line templates $A_{em}$ Four emission lines that are in WFC3/G141 bandpass for MRG-S0851, are included. They are: [O III], H $\beta$ , H $\gamma$ , and H $\delta$
$x_c$	$[A_{c,l}]_X$	The weight $x_c$ for the contamination $A_c$ of MRG-S0851 by other objects.
$[x_{b,k}]_G$	$[A_{b,kl}]_{G \times X}$	The different exposures could have different constant backgrounds. The constants are $x_b$ (“bias” term in regression), and the background model is $A_b$ . Here, $G$ is the number of exposures.
$[x_{p,in}]_{M \times I}$	$[A_{p,inl}]_{M \times I \times X}$	The weight $x_p$ of the polynomial fit $A_p$ to the data. $I$ is the degree of the polynomial used.

**Note.** All matrices are denoted with the brackets and their shapes are shown as indices. Note that, to reduce computation cost, we eventually turn these matrices to 1D arrays (see Figure 4). The first and second columns indicate the weight, and its corresponding “predictor” is followed by a description in the third column.

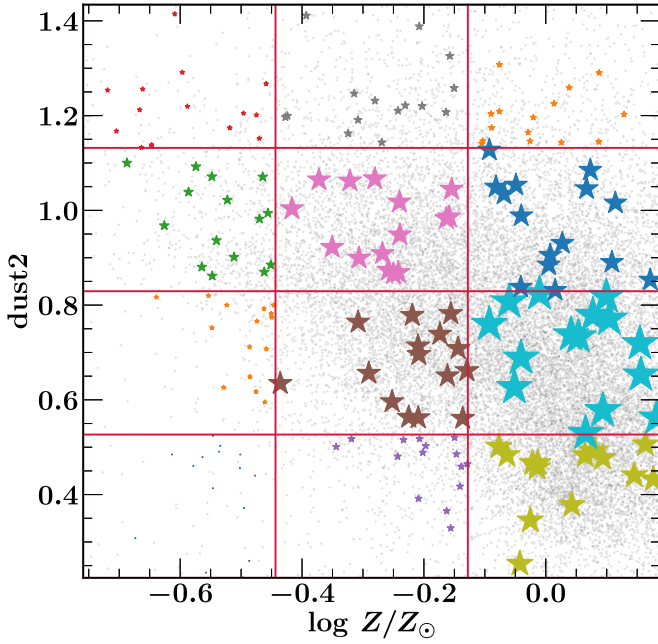
#### 4.3.1. The Building Blocks of the Linear Regression Model

Modeling a CSP using a linear combination of SSP templates is a generalized linear regression problem, the elements of which are shown in Table 2. We use the FSPS models and its Python wrapper (Foreman-Mackey et al. 2014), assuming the dust and metallicity values from the Prospector- $\alpha$  posterior (Section 4.2). The SSP spectra vary in age, starting at 10 Myr and ranging up to the age of the universe at the redshift of the main science target, increasing on a logarithmic scale, such that the logarithm of the ratio of two adjacent ages in Gyr is 0.05, i.e.,  $\log t_{i+1}[\text{Gyr}]/t_i[\text{Gyr}] = 0.05$ .<sup>19</sup> These spectral models are then used to simulate the corresponding 2D G141 grism spectra using Grizli. We also add a first-degree polynomial to the FSPS templates. By fitting a polynomial to grism spectra, we address any issues in the background such as enhanced airglow for a particular orientation and/or contamination (Brammer et al. 2012a). In a joint spectrophotometric fit in particular, this polynomial fit addresses any spectroscopic flux calibration

errors and tunes the spectral continuum shape to the photometry (Newman et al. 2018a), noting that the photometric data is solely modeled by SSPs with no extra polynomial being fitted. This prevents the continuum shape of the grism spectrum from solely dictating the dust solution, as this should be mostly determined by the longer wavelength baseline of photometry. Also, by keeping the polynomial degree to lower values (usually less than 3, and for our pilot target, MRG-S0851, we pick 1), we prevent it from generating any spikes or spectral features that could affect age/metallicity measurements.

In order to use the Prospector- $\alpha$  posteriors as the dust and metallicity priors to generate SSPs for resolved fitting with requiem2d, we project the full posterior into the dust2 and  $\log Z/Z_\odot$  plane, limiting the extension of each axis of this plane to the  $3\sigma$  width of the corresponding credible interval. We note that our Prospector- $\alpha$  fit assumes the Kriek & Conroy (2013) dust model, which has three free parameters, including dust2, which controls the attenuation of stellar populations older than  $10^7$  yr; for further discussion, see, e.g., Noll et al. (2009), Conroy & Gunn (2010), Kriek & Conroy (2013), and our Section 6.1. We then define  $3 \times 4$  boxes in this plane, drawing 15 samples from the full Prospector- $\alpha$  posterior in each box, calculating the median of the draws. In other words, we use the 2D projection in  $\log Z/Z_\odot$ -dust2 plane to draw samples from the full Prospector- $\alpha$  posterior of dust and metallicity, which has four parameters. Figure 3 shows the

<sup>19</sup> In practice, we generate a series of nonoverlapping constant SFHs that include each age in our grid at their center. This is argued to be a more realistic approximation than pure SSPs (e.g., see Morishita et al. 2019). We test both cases, but we do not find any significant difference, potentially owing to the finer sampling of the age grid in our study, i.e.,  $\sim 50$  compared to the 10 of Morishita et al. (2019).



**Figure 3.** Prospector- $\alpha$  posteriors of dust2 and metallicity for MRG-S0851 are shown by light gray points. We show different boxes, defined to sample the posterior to generate SSPs, with red lines. Actual draws in each box are indicated by stars. Size of stars demonstrate the weight of each box in the Prospector- $\alpha$  posterior.

2D projection of the Prospector- $\alpha$  posterior for the global analysis of the pilot target, MRG-S0851.

We have 12 sets of templates, each corresponding to a different region in dust and metallicity. Each one of the 12 sets of templates has a weight that is inferred by summing the weights of individual draws from the Prospector- $\alpha$  posterior falling into the corresponding box posterior and the selections described here (Figure 3). We rank-order the SSPs using the final weights and sample the weights using a stick-breaking Dirichlet process (e.g., Connor & Mosimann 1969; Sethuraman 1994) with  $\beta \sim \text{Beta}(1, \alpha)$  and  $\alpha \sim \text{Gamma}(11, 1)$ . In this process, one draws a set of initial 12 weights,  $\beta'_i, i = 1 \dots 12$  from the beta distribution. Values of  $\beta'_i$ s are between 0 and 1, but they do not necessarily add up to unity; to make sure that they do, the final set of weights is calculated using  $\beta_i = \beta'_i \Pi_{j=1}^{i-1} (1 - \beta'_j)$ , analogous to breaking a stick of length 1.

For our pilot target, MRG-S0851, we have four major emission lines filling the underlying absorption features in G141 bandpass: H $\beta$ , H $\gamma$ , H $\delta$ , and [O III]. We include a separate template for each one of these emission lines from Grizli: A Gaussian one-dimensional spectral template centered at the wavelength of each emission line is normalized to unity and is convolved with MRG-S0851 morphology to generate a two-dimensional grism template. The coefficients of these templates are being fitted with the rest of parameters using the Monte Carlo method, providing an estimate on the strength of emission lines.

We multiply all photometric bands in the model of each spatial bin with a set of nuisance parameters  $\omega$ , with a prior of  $N(1, 1)$ , i.e., a normal distribution with  $\mu = 1$  and  $\sigma = 1$ , to address any calibration mismatch between the photometric and spectroscopic data. The general model for one set of SSPs with all elements in place can then be described by the following

equations (see Table 2 for the description of each element):

$$M_{s,l} = \sum_{i=1}^M \sum_{n=1}^I x_{p,in} A_{p,inl} + \sum_{i=1}^M \sum_{j=1}^N x_{ij} A_{s,ijl} + \sum_{i=1}^M \sum_{q=1}^4 x_{em,iq} A_{em,iq} + x_c A_{c,l} + \sum_{k=1}^G x_{b,k} A_{b,kl} \quad (1)$$

$$M_{p,ir} = \omega_i \sum_{j=1}^N x_{ij} A_{p,ijr}, \quad (2)$$

where  $M_{s,l}$  denotes the 2D grism model of the  $l$ th HST pixel, and  $M_{p,ir}$  indicates the photometric model for the  $i$ th spatial bin and the  $r$ th photometric band.

#### 4.4. The Priors and the Monte Carlo Sampling Method in requiem2d

We determine the posterior distribution of weights in the generalized linear regression model (defined in Equations (1) and (2)) using Bayes' theorem. The photometry likelihoods are assumed to follow a mixture of normal distributions with a standard deviation estimated from the observational errors and weights from the Dirichlet process. To be more specific, for each one of 12 sets of SSPs (Section 4.3.1), we generate a model using the SFH model and calculate the photometric fluxes. We next assume 12 normal distributions centered at these fluxes with standard deviations equal to the observed photometric uncertainty. The full likelihood probability distribution is then the weighted sum of the 12 normal distributions with weights determined through a stick-breaking Dirichlet process, as described in Section 4.3.1.

As we have more than 1000 grism pixels for each spatial bin in spatially resolved spectroscopic data, we adopt a simplifying assumption that the final grism model is a weighted average of 12 SSPs. To maintain consistency between the resolved and unresolved analysis, we apply the same assumption to the spatially unresolved spectroscopic data. We test this simplifying assumption explicitly for the unresolved analysis of MRG-S0851 by sampling the age posterior twice, first using a full mixture of normal distributions for grism spectroscopy and then using the weighted average of 12 SSPs. No statistically significant difference is detected in recovered ages adopting these two approaches.

The prior of weights,  $\mathbf{x}$ , is derived from the SFH prior. As mentioned in Section 4.2.1, we adopt a continuity prior for the SFH following a regularizing scheme introduced for the same problem in Ocvirk et al. (2006):

$$\log \text{SFR}_{n,t} - 2 \log \text{SFR}_{n,t-1} + \log \text{SFR}_{n,t-2} \sim \epsilon_t, \quad (3)$$

with  $\epsilon_t \sim N(0, 1/20)$ . For a linearly defined age grid, this can also be interpreted physically as a requirement of the continuity for the first time derivative of SFR (the slope of SFR, or SFR increments). Other versions of a continuity prior may also be used. For example, Leja et al. (2019a) require the continuity of the SFR itself.<sup>20</sup> We find that, in our case, analyzing massive quiescent galaxies at  $z \sim 2$ , the continuity of the SFR slope recovers SFHs and ages slightly better than the continuity of

<sup>20</sup> We note that Leja et al. (2019a) adopt Student's  $t$ -distribution for  $\epsilon_t$  in the right-hand side of Equation (3).



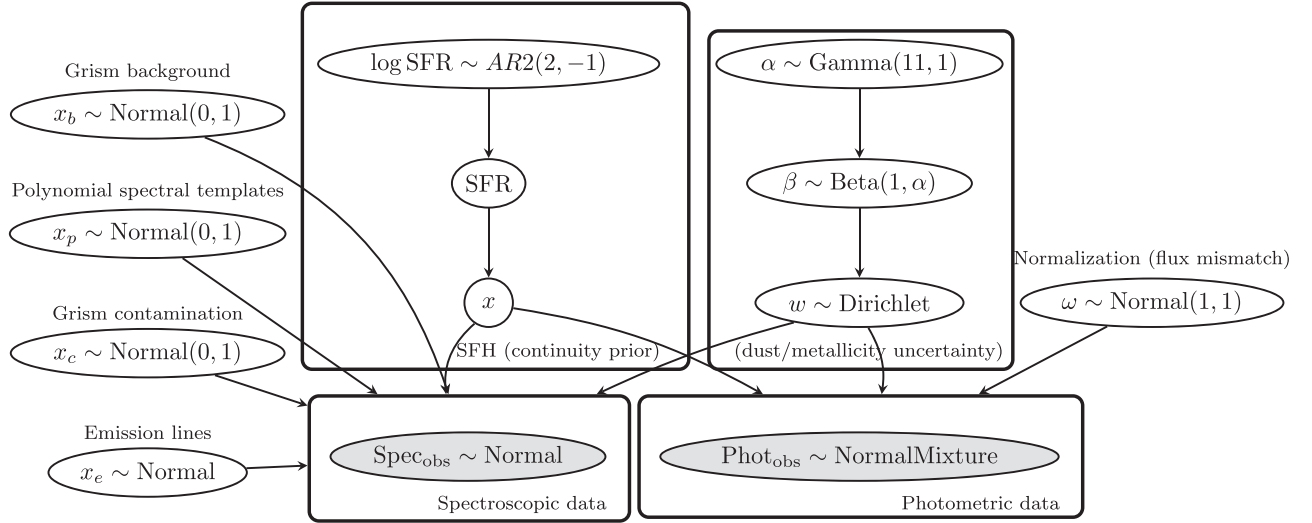


Figure 4. Statistical model for the spatially resolved analysis demonstrated using the plate notation.

the SFR. This may be because it is a stronger prior requirement, helping with the finer sampling of SSPs at lookback times greater than  $\sim 1$  Gyr.

To connect SFRs to the weights, each SSP has a mass-to-light ratio which we use to calculate the corresponding mass weight,  $x^M$ , from  $x$ . This mass weight can then be connected to SFR using  $\text{SFR}_t = x_t^M / \delta t$ . The rest of the priors and the model itself are demonstrated in Figure 4.

Estimating the age, weights  $x$ , or other parameters of the stellar populations from the observed spectra could be ill-posed, and doing so usually requires regularizing (Ocvirk et al. 2006). Also, due to highly correlated parameters and the higher dimension of the problem, the usual Monte Carlo algorithms such as random-walk Metropolis (Metropolis et al. 1953) fail to sample the posterior efficiently (Neal 1993). We therefore use No-U-Turn Sampling (NUTS; Homan & Gelman 2014), which is a variation of the Hamiltonian Monte Carlo method (Neal 2011) to sample the posterior. NUTS uses a recursive algorithm to build a set of points spanning a wide swath of the target distribution, stopping automatically when it starts to retrace its step (Homan & Gelman 2014). NUTS is proven to be more efficient in exploring correlated parameter spaces, such as in our problem, relative to the random-walk methods (Creutz 1988; Neal 2011; Homan & Gelman 2014).

We use the Python package (pymc3; Salvatier et al. 2016) extensively in our analyses for sampling the posteriors with NUTS. Two chains are constructed, drawing 1000 (unresolved analysis) and 1400 (resolved analysis) samples in each one, considering only the second half of the chains as post burn-in draws. We check for divergences using Gelman–Rubin statistics (Gelman & Rubin 1992) explicitly and combine the chains.

## 5. Inferring the Age of Stellar Populations

In this section, we show how we can infer the age of stellar populations and its uncertainty from the weights,  $x$ , defined in Equations (1) and (2).

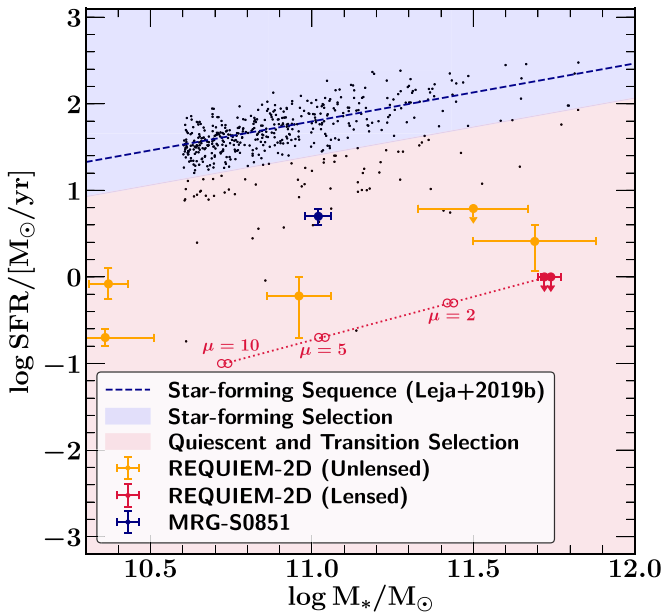
The posterior of weights,  $x$ , which is the basic output of *requiem2d*, can be interpreted as the light weight of each SSP template. We can use these weights to calculate the light-weighted average ages; however, light-weighted ages are

misleadingly young, as younger stars outshine the older stars. We therefore use the mass-to-light ratio of SSP templates to calculate the mass weights,  $x^M$ . This quantity is used to reconstruct SFHs and to calculate the median mass-weighted age  $t_{50}$ ,<sup>21</sup> shown to be a robustly estimated from models (e.g., Belli et al. 2019). The uncertainty of the median mass-weighted age is estimated directly from the Monte Carlo chains. The median mass-weighted age,  $t_{50}$ , is also independent of the lensing magnification, as any effect of magnification on SFR is canceled out in the definition of  $t_{50}$  (like sSFR).

The final goal of the *requiem2d* code is to recover both global and resolved ages and SFHs of massive quiescent galaxies, and as it uses a nonparametric SFH history in a Bayesian framework, it is important to choose an appropriate SFH prior. We test our methodology and our choice of prior, Equation (3), using a sample of massive galaxies with  $\log M_*/M_\odot \geq 10.6$  at  $z = 2$ , selected from Illustris, a cosmological hydrodynamical simulation of galaxy formation with a volume of  $(100 \text{ Mpc})^3$  that includes a comprehensive physical model (Genel et al. 2014; Vogelsberger et al. 2014a, 2014b; Nelson et al. 2015). To obtain the Illustris SFHs for our test, we construct the histograms of the formation times of all the star particles in the galaxy, weighted by the masses of the star particles at  $z = 2$ . The Illustris SFHs should therefore be directly comparable to our mass-weighted SFHs. Using Illustris SFHs, we perform three tests: we test if we can recover (1) global SFHs and ages of quiescent and transition galaxies, (2) age gradients of quiescent galaxies, and (3) global ages and SFHs of star-forming galaxies. We will discuss these tests each in turn.

A subsample of massive galaxies with low sSFRs is selected by requiring that the SFR at the closest snapshot at  $z = 2$ , corresponding to the average SFR in the last  $\sim 30$  Myr given the width of our time bins, is 0.4 dex below the empirical star-forming sequence in the  $\log(\text{SFR})$ – $\log(\text{stellar mass})$  plane of Leja et al. (2019b), noting that the locus of the star-forming main sequence in Leja et al. (2019b) is 0.3 dex lower than that of Whitaker et al. (2014a). Our Illustris selection consists of a more diverse sample than just traditional quiescent galaxies, including

<sup>21</sup> The formal definition of  $t_{50}$  is  $\int_{t_0}^{t_0} dt' \text{SFR}(t') = 0.5 \times \int_0^{t_0} dt' \text{SFR}(t')$ , where  $t_0$  is the age of the universe at the redshift of interest.



**Figure 5.** Sample of massive galaxies in the  $\log(\text{SFR})$ – $\log(\text{stellar mass})$  plane, selected from the Illustris cosmological simulation to test our methodology and the choice of SFH prior. Black dots are the Illustris galaxies, with our REQUIEM sample indicated by larger circles using global SFRs measured by Newman et al. (2018a), A. Man et al. (2020, in preparation), and this work. SFR and stellar mass of two targets, MRG-S1522 and MRG-P0918, are not demagnified; we plot their projected locations assuming magnifications of  $\mu = 2, 5, 10$  (open circles) along the dotted-line trajectory of increasing magnification. Color-coding identifies our subsamples of quiescent and transition (red) and star-forming populations (blue).

galaxies below the main sequence but above the quiescent population according to Leja et al. (2019b). Therefore, we call this population “quiescent and transition galaxies” hereafter. This selection leaves room for new discoveries and makes our test more robust. The region in  $\log(\text{SFR})$ – $\log(\text{stellar mass})$  parameter space corresponding to our combined quiescent and transition selection is shown in Figure 5. Out of 502 galaxies initially selected from Illustris, given our stellar mass cut of  $\log M_*/M_\odot \geq 10.6$  at  $z = 2$ , 71 galaxies are included in our quiescent and transition subsample. The stellar masses of the subsample have a distribution of  $\log M_*/M_\odot = 11.03^{+0.30}_{-0.37}$ , where the upper and lower limits correspond to the 84th–50th and 50th–16th percentiles. Figure 6 (left panel) shows the final SFHs of quiescent galaxies selected from Illustris, rank-ordered by their median ages.

We generate spectral templates using FSPS models from the Illustris SFHs, fixing the dust and metallicity to the best-fit values of MRG-S0851 (to be presented in Section 6.1). The morphology of MRG-S0851 is used to generate a mock grism spectrum, and we choose the same five HST and two Spitzer photometric bands in order to generate simulated photometric data. Noise is added to grism pixels by assuming an S/N that is drawn randomly for each target from a uniform distribution of  $U(0.01, 0.1)$  for grism pixels. We assume that the S/Ns of photometric measurements are five times higher than the S/Ns of grism pixels, in the range of 50–500. For comparison, the S/N of 12-orbit depth grism data of MRG-S0851 is  $\sim 10$ –20 (corresponding to noise at the 5%–10% level) and the photometric S/N is  $\sim 20$ –200 with  $< 1$  orbit of imaging data.

The results of the simulations are presented in Figures 6 and 7. Figure 6 demonstrates that the general average trends of SFHs are recovered reasonably well. However, one shortcoming of this methodology is that it struggles to recover particularly stochastic SFHs (see the right column, Figure 6). This is potentially due to the continuity prior disfavoring stochastic jumps. Figure 7 shows recovered median ages versus mock ages from Illustris (left panel) and its scatter as a function of grism noise per pixel. There are no noticeable systematic biases in the recovered ages, except for ages older than  $\sim 1.5$  Gyr that appear to be slightly younger (Figure 7). We note that having a noisier grism spectrum increases the scatter around the recovered age, and by increasing the level of uncertainty from 0.01 dex to 0.05 dex per grism pixel, the scatter of median ages increases from  $\sim 0.02$  to  $\sim 0.1$  dex (Figure 7, right panel). The results of these mock tests are therefore useful for planning future grism observations, where Figure 7 gives the required S/N for a given target age accuracy.

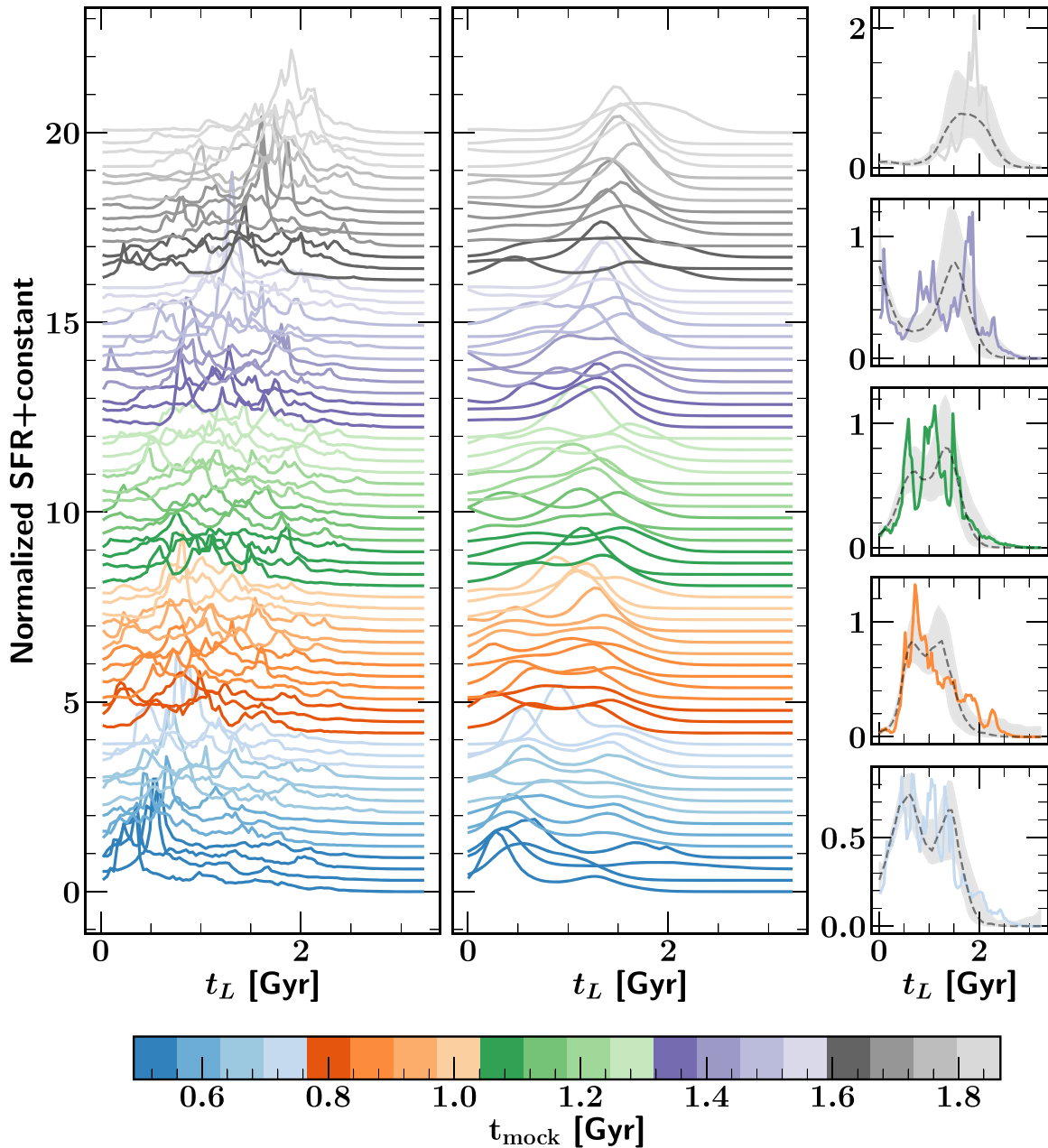
Next, we test the age gradients. To generate the mock observations, we use the MRG-S0851 morphology, defining three spatial bins with widths of  $0''.18$  at the center of MRG-S0851. We assume that two adjacent bins have the same SFHs and ages (perfect symmetrical galaxy). We randomly select an SFH from the subsample of Illustris quiescent and transition SFHs, selected as described in the first test, for the central bin. For adjacent bins, we then randomly select another Illustris quiescent and transition SFH such that the age gradient between the center and the two adjacent bins lies within  $\pm 0.5$  Gyr. Next, we generate FSPS models as described in the first test. Finally, we use Grizli to simulate the grism spectrum of three bins. To be fully consistent with the real data, five resolved HST photometric bands for each one of the three bins are calculated, but only global/unresolved Spitzer IRAC channels 1 and 2 photometric measurements are assumed. We generate 50 mock observations and add noise in the exact same way as the first test by assuming a constant S/N per grism pixel. However, we note that, in this test, we are only covering a portion of the galaxy, and the total flux associated with each SFH is relatively lower than the first test.

Figure 8 shows the age gradient test results, where we demonstrate the recovered  $\Delta t = t_{50, \text{center}} - t_{50, \text{outskirt}}$ . There are no systematic biases in the recovered age gradients; however, we caution that having a sharper gradient seems to increase the scatter.

As for the third test, we randomly select 100 galaxies from the complement of the quiescent and transition region in the  $\log(\text{SFR})$ – $\log(\text{stellar mass})$  plane, i.e., we select 100 galaxies whose SFR is greater than 0.4 dex below the empirical star-forming sequence of Leja et al. (2019b) (blue region in Figure 5). The mock data are generated following the exact same steps as were taken in the first two tests.

Figure 9 shows an aggregate plot of the SFHs of massive star-forming galaxies at  $z \sim 2$  selected from the Illustris simulation. The `requiem2d` code and the continuity prior can recover the general trends of SFHs reasonably well—with the notable exception of some difficulty recovering highly stochastic features in the SFHs, as was the case with analyses of massive quiescent galaxies (see Figure 6).

Figure 10 shows the recovered median mass-weighted ages versus median mock ages from massive star-forming galaxies in the Illustris simulation (left panel), as well as the scatter of



**Figure 6.** Testing the *requiem2d* methodology using SFHs of massive quiescent and transition galaxies at  $z = 2$  selected from the Illustris simulation. Figure shows Illustris SFHs and recovered SFHs using *requiem2d* vs. lookback time,  $t_L$ . SFHs are color-coded by their mock Illustris median ages,  $t_{\text{mock},50}$ . Illustris SFHs (left panel) agree well with the recovered SFHs (middle panel). Right panel shows five SFHs with a range of ages and their recovered models in more detail. Dashed gray line and gray shading indicate the median and  $1\sigma$  width of recovered SFH using *requiem2d*. Second galaxy from the top in the right panel has an SFH that looks similar to what we recover for MRG-S0851 in the last  $\sim 1$  Gyr of evolution. Top right panel shows the oldest galaxy in the sample with a median age of 1.9 Gyr.

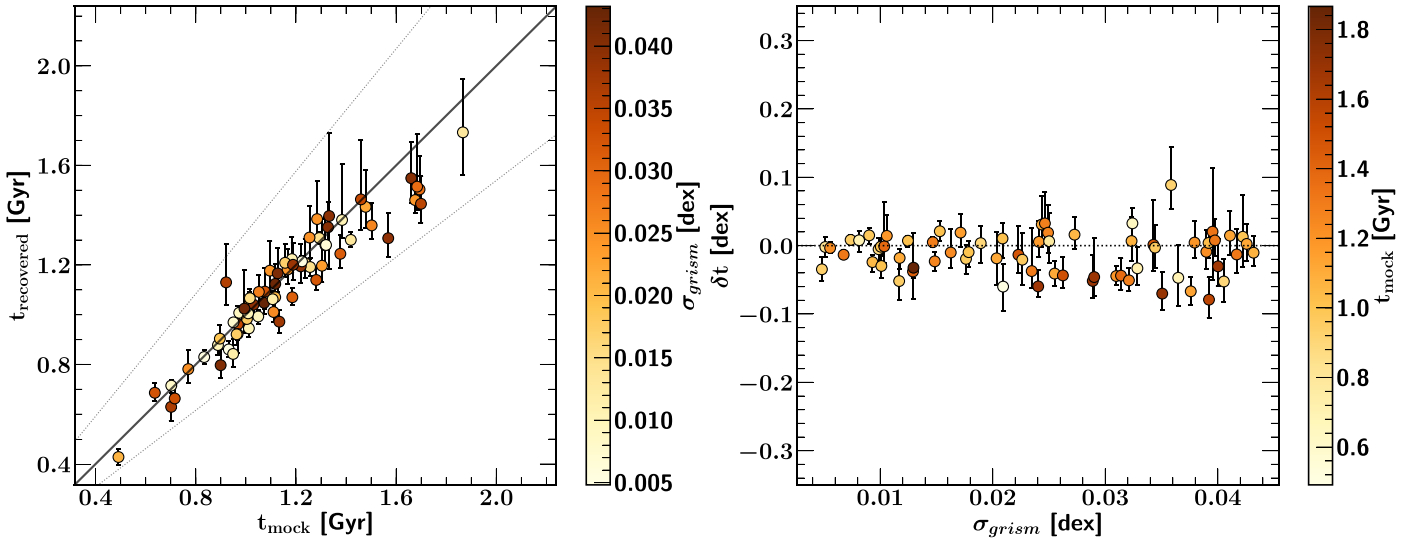
median ages versus the grism pixel noises (right panel). The recovered ages match reasonably well with the mock ages. We also notice no systematic offsets or biases when considering the recovered ages in Figure 10. These results demonstrate that the *requiem2d* code, with the continuity prior (Equation (3)), can also be used to derive robust constraints on the ages of massive star-forming galaxies at  $z \sim 2$ .

To conclude, we caution that a careful choice of SFH prior is important in nonparametric SFH models (e.g., Leja et al. 2019a). Our choice of the continuity of the SFR slope, Equation (3), seems to be appropriate for analyzing the SFHs and ages of massive galaxies at  $z \sim 2$ , as justified by recovering ages and SFHs from the Illustris cosmological simulation.

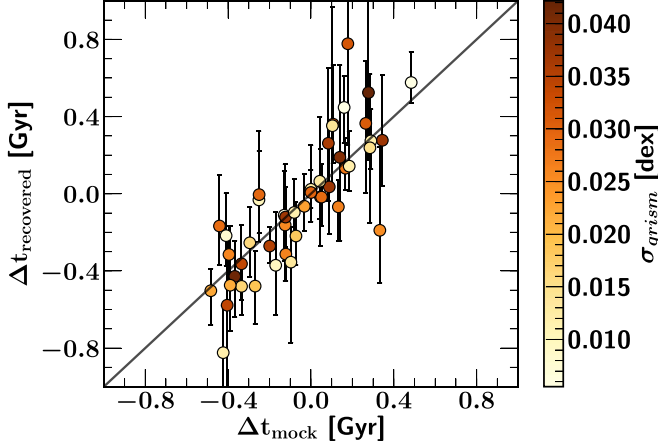
## 6. Testing Our Methodology with Real Observations: A Pilot Study of MRG-S0851

SDSS J0851+3331-E was first detected in the Sloan Giant Arcs Survey (SGAS), which is a survey of strongly lensed galaxies (Hennawi et al. 2008; Bayliss et al. 2011; Sharon et al. 2020). The cluster is at a redshift of  $0.3689 \pm 0.0007$  with a R. A. of 8:51:39 and a decl. of +33:31:10.83; see Tables 1–3 and Section 4.3.6 in Sharon et al. (2020) for more detailed information and discussion. Throughout this paper, our target is referred to as MRG-S0851, as it is a Magnified Red Galaxy (MRG, following Newman et al. (2018a)) and represents a pilot analysis for the REQUIEM survey.





**Figure 7.** Testing our methodology to recover global SFHs and ages using a sample of massive quiescent and transition galaxies selected from the Illustris simulation. Left panel: recovered vs. actual global  $t_{50}$  ages, color-coded by grism noise. Solid line is one-to-one relation. Dotted lines are  $\pm 0.1$  dex scatter. Right panel: deviation of median recovered  $t_{50}$  ages from the true ages vs. grism noise, color-coded by true ages. No noticeable systematic biases can be seen, indicating that our choice of prior for SFHs (Equation (3)) reasonably recovers the ages of massive quiescent and transition galaxies at  $z = 2$ , selected from the Illustris simulation.



**Figure 8.** Recovering age gradients of quiescent and transition galaxies using our methodology for three spatial bins with widths of  $0''.18$  that have five resolved HST photometric measurements, two unresolved Spitzer IRAC channels 1 and 2 measurements, and grism spectra. SFHs are selected from the Illustris simulation.

MRG-S0851 was observed with HST (program HST-GO-13003, PI: M. Gladders) (Sharon et al. 2020), as well as the Spitzer Space Telescope (Rigby et al. 2012). The Spitzer data for MRG-S0851 include observations with the IRAC channels 1 and 2 on 2010 December 13, as a part of a larger campaign of infrared imaging for the SGAS survey (Spitzer proposal IDs 90232, PI: J. Rigby, and 70154, PI: M. Gladders). The HST/WFC3 observations include the IR filters  $H_{F160W}$  and  $J_{F125W}$ , as well as the UVIS filters  $I_{F814W}$  and  $U_{F390W}$  on 2016 February 26.

We combine the existing observational data with our HST program for MRG-S0851 (HST-GO-14622, PI: K. Whitaker), adding the WFC3 IR filters  $H_{F160W}$ ,  $Y_{F105W}$  and 12 orbits with the G141 grism dispersing element. For this follow-up program, each of the 12 orbits contained a short imaging exposure with either  $H_{F160W}$  or  $Y_{F105W}$  ( $\approx 250$  s of exposure time) followed by a longer exposure with the grism G141 ( $\approx 2400$  s of exposure time). These 12 orbits of data were executed between 2017 January 28 and 2017 February 15. To

minimize the contamination of the MRG-S0851 grism spectra from nearby bright cluster members, we used two different dispersion angles of  $P_\theta \approx 29^\circ$  and  $P_\theta \approx 35^\circ$  in the G141 observations.

Figure 1 shows the final data product for MRG-S0851 for  $H_{F160W}$  and the G141 grism. In the left panel, we show the  $H_{F160W}$  mosaic, the deepest photometric data. In the right panel, we show the drizzled image of WFC3/G141 grism spectra for six orbits, corresponding to  $P_\theta \approx 29^\circ$  orientation angles. The 2D grism spectrum of the brightest image, E3, is shown in Figure 2.

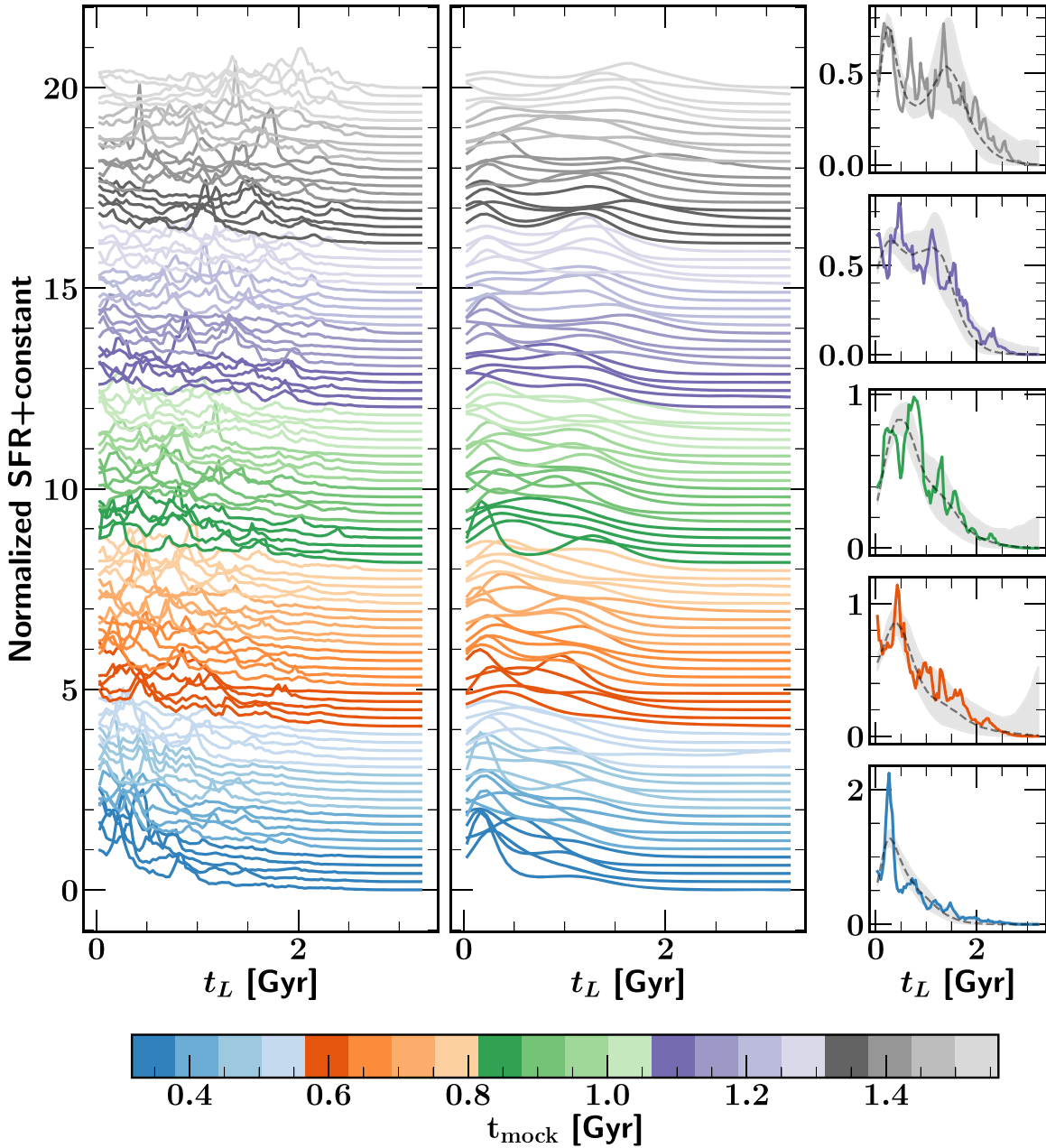
The details of the lensing model and the morphological analysis of MRG-S0851 are discussed in Appendices A and B, respectively. While MRG-S0851 is quintuply imaged, we will concentrate on image E3 in this paper, since other images are only partial according to our lens model, and their light profiles do not represent the full light profile of MRG-S0851 (see Figure B2).

### 6.1. Results of the Unresolved Photometric Analysis

The unresolved global stellar population fit to seven photometric bands for MRG-S0851, obtained using *Prospector- $\alpha$*  as discussed in Section 4.2, is shown in Figure 11. The *Prospector- $\alpha$*  fit yields a mass-weighted average age of  $2.0 \pm 0.3$  Gyr for the unresolved stellar populations, and an sSFR of  $\log \text{sSFR}_{100 \text{ Myr}} / [\text{yr}^{-1}] = -9.8^{+0.3}_{-0.2}$ , independent of the lensing magnification.

The demagnified stellar mass of MRG-S0851 is estimated to be  $\log M_*/M_\odot = 10.96 \pm 0.04$  and the demagnified SFR is constrained to be  $\log \text{SFR}_{100 \text{ Myr}} / [M_\odot \text{ yr}^{-1}] = 1.2^{+0.3}_{-0.2}$ , correcting for strong gravitational lensing magnification of  $\mu_{E3} = 5.7^{+0.4}_{-0.2}$ . The propagated uncertainty of gravitational magnification on these parameters is negligible relative to the uncertainty of the photometric fit. As discussed in Section 4.2.1, these results are obtained using photometric data alone, and we next jointly fit spectrophotometric data to obtain constraints on the age of the stellar populations.

The dust and metallicity values for MRG-S0851 are reported in Table 3. The best-fit *dust2* parameter of MRG-S0851 implies an  $A_V$  of  $\sim 0.8$ ; however, as *dust\_index*  $\sim 0.15$ , the



**Figure 9.** Testing the `requiem2d` methodology using SFHs of massive star-forming galaxies at  $z = 2$  selected from the Illustris simulation. Panels and labels are exactly the same as in Figure 6, but for massive star-forming galaxies.

attenuation curve will be flatter than the Calzetti law, and we have comparatively less UV attenuation and more near-IR attenuation, as discussed in Section 4.2.1. The measured metallicity value has a large uncertainty, consistent with both a subsolar and a solar metallicity within the error bars. A tighter constraint is necessary in order to speculate about the future chemical enrichment of MRG-S0851.

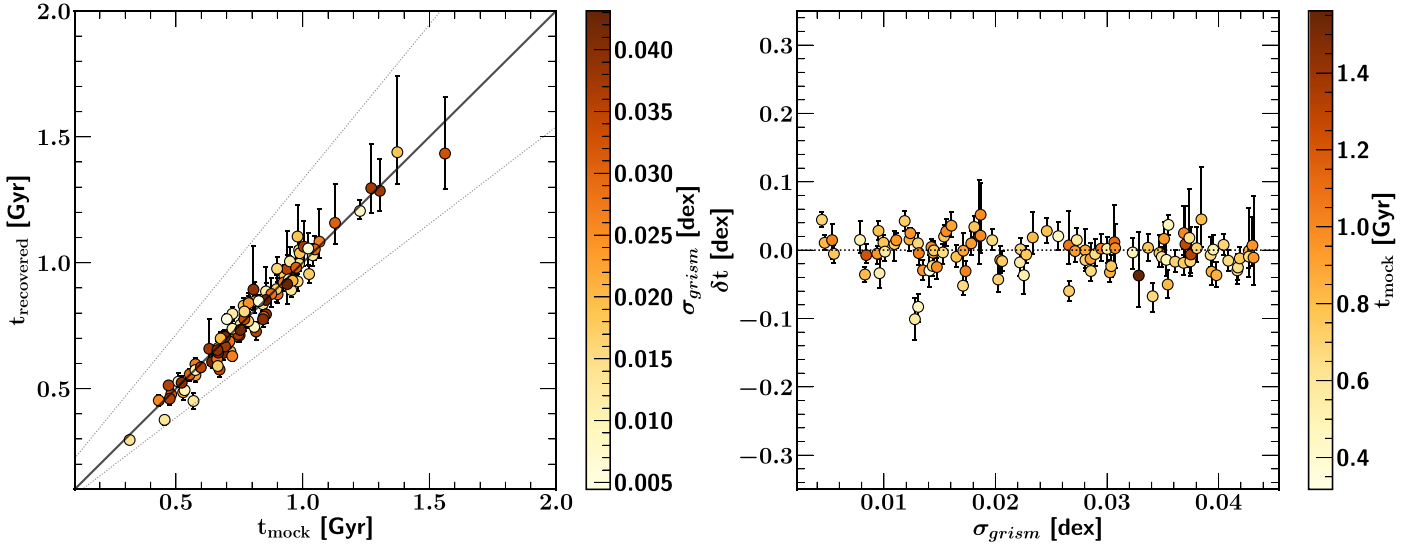
### 6.2. Results of the Spectrophotometric Analysis for the Case Study of MRG-S0851

In this section, we apply the methodology described in Section 4 to MRG-S0851. We include the photometric data obtained from five WFC3 filter images ( $H_{F160W}$ ,  $J_{F125W}$ ,  $Y_{F105W}$ ,  $I_{F814W}$  and  $U_{F390W}$ ) along with the spectroscopic data of WFC3-G141 in our joint fitting. For the unresolved analysis, we explicitly

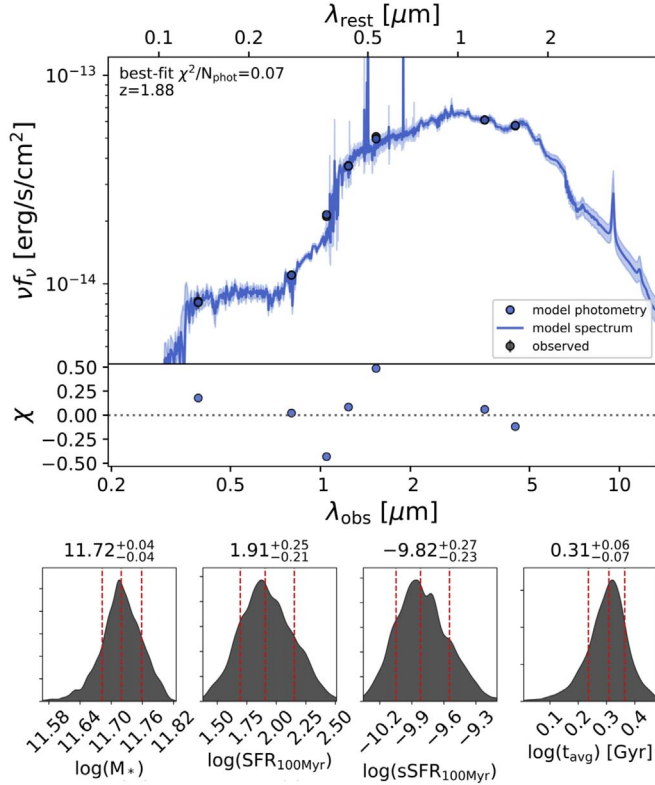
include the two Spitzer IRAC bands, and for the resolved analysis, we require that the sum of spatial bins' IRAC fluxes matches two unresolved IRAC fluxes as an additional constraint.

Figure 12 shows the seven spatial bins for each grism exposure obtained following the recipe of Section 4.1.1. We explicitly check that the center bin also includes the peak of the  $U_{F390W}$  light profile.

Each pixel in the image is  $0''.06$ , so the central bins range from  $0''.18$  to  $0''.24$  wide. Using our lensing model (see Appendix A), the bins are centered at distances of  $<-1.4$  kpc,  $-1.0$  kpc,  $-0.5$  kpc,  $0$ ,  $+0.8$  kpc,  $+1.2$  kpc, and  $>+1.5$  kpc. The five central bins therefore span a total of 2.9 kpc in the source plane, which is comparable to the half-light diameter of  $\simeq 3.4$  kpc measured in Appendix B.1. We therefore probe the age gradient of the stellar populations in the inner radius of  $\simeq 1.7$  kpc of MRG-S0851 at an average spatial resolution of  $\simeq 0.6$  kpc.



**Figure 10.** Testing our methodology to recover global SFHs and ages using a sample of massive star-forming galaxies selected from the Illustris simulation. Left panel: recovered vs. the actual global  $t_{50}$  ages, color-coded by grism noise. Right panel: deviation of median recovered  $t_{50}$  ages from the true ages vs. grism noise, color-coded by true ages. Symbols and the labels are exactly the same as in Figure 7, but for massive star-forming galaxies.



**Figure 11.** Unresolved global stellar populations fit, which is obtained using *Prospector- $\alpha$* , to five HST bands and two Spitzer IRAC channels. Stellar mass and star formation rate (SFR) are not corrected for strong gravitational lensing effect in this plot.

The results for the unresolved analysis of the stellar populations of MRG-S0851 are shown in Figure 13, with the 1D grism spectra in the right panel, extracted following Horne (1986).<sup>22</sup> The spectral energy distribution (SED) of the unresolved stellar populations can be found in the left panel of

**Table 3**

The Best-fit Values of Dust and Metallicity for the Global Photometric Measurements of MRG-S0851 (5 HST + 2 Spitzer Bands) as Well as Spatially Resolved Bins (5 HST Bands Only), as Defined in Section 4.1.1 and Shown in Figure 12, Obtained by *Prospector- $\alpha$*

	dust1	dust2	dust_index	log $Z/Z_{\odot}$
Global	$0.69^{+0.72}_{-0.43}$	$0.73^{+0.43}_{-0.39}$	$0.15^{+0.22}_{-0.38}$	$-0.04^{+0.21}_{-0.44}$
$< -1.4$ kpc	$0.66^{+0.71}_{-0.41}$	$0.87^{+1.06}_{-0.65}$	$-0.57^{+0.46}_{-0.38}$	$-0.24^{+0.40}_{-0.84}$
$-1.0$ kpc	$0.83^{+1.34}_{-0.64}$	$0.87^{+1.06}_{-0.65}$	$0.05^{+0.29}_{-0.54}$	$-0.17^{+0.33}_{-0.52}$
$-0.5$ kpc	$0.9^{+1.44}_{-0.78}$	$0.84^{+1.07}_{-0.68}$	$0.15^{+0.22}_{-0.75}$	$-0.14^{+0.3}_{-0.49}$
Center	$1.25^{+1.46}_{-1.18}$	$1.22^{+1.05}_{-1.15}$	$0.32^{+0.08}_{-1.05}$	$-0.11^{+0.28}_{-0.42}$
$+0.8$ kpc	$1.02^{+1.53}_{-0.95}$	$1.02^{+1.14}_{-0.93}$	$0.24^{+0.15}_{-0.58}$	$-0.17^{+0.33}_{-0.36}$
$+1.2$ kpc	$0.93^{+1.28}_{-0.73}$	$0.96^{+0.98}_{-0.72}$	$-0.04^{+0.39}_{-0.76}$	$-0.15^{+0.32}_{-0.59}$
$> +1.5$ kpc	$0.71^{+1.3}_{-0.6}$	$0.69^{+1.03}_{-0.57}$	$-0.28^{+0.62}_{-0.66}$	$-0.08^{+0.24}_{-0.52}$

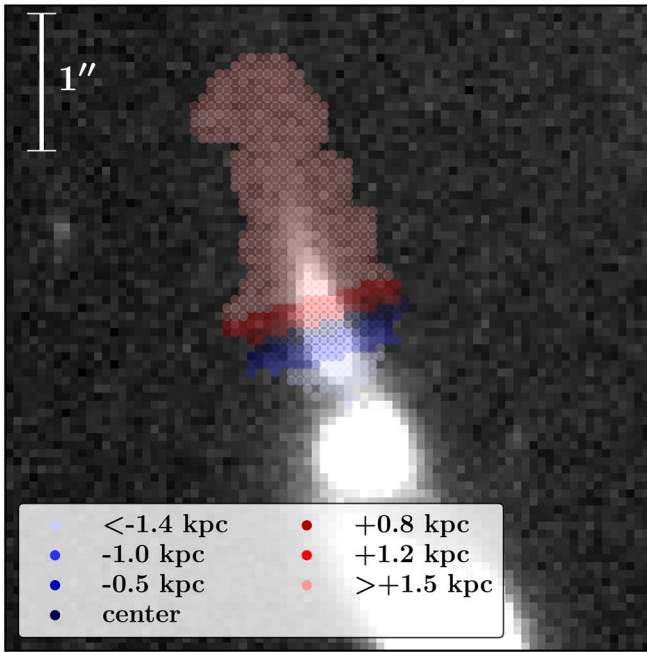
**Note.** Kriek & Conroy (2013) dust model is assumed here. In this model, dust1 and dust2 are the parameters controlling the strength of dust attenuation for the stellar populations younger and older than  $10^7$  yr, respectively, and dust\_index controls the slope of the dust attenuation curve.

Figure 13. MRG-S0851 has significant flux in the rest-frame UV, which is sampled by the photometric filter  $U_{F390W}$ . This rest-frame UV flux indicates a recent star formation activity detected in the last  $\sim 100$  Myr of evolution in MRG-S0851, which is overall  $\sim 1.4$  dex below the star-forming sequence of Whitaker et al. (2014a) and  $\sim 1.1$  dex below the star-forming sequence of Leja et al. (2019b) (based on the *requiem2d* joint fit). We will discuss this result in depth in a subsequent paper (M. Akhshik et al. 2020, in preparation).

We measure a global median mass-weighted age of  $1.8^{+0.3}_{-0.2}$  Gyr with a corresponding global light-weighted age of  $1.4^{+0.2}_{-0.1}$  Gyr from the joint fitting. The results for the five central bins of MRG-S0851 are also shown in Figure 14. Figure 15 shows the median mass-weighted age for both the resolved and unresolved stellar

<sup>22</sup> The 1D extraction is only performed to make the final plot at the end of the analysis; we fit and analyze the grism spectrum in its native 2D space.





**Figure 12.**  $H_{F160W}$  image of MRG-S0851, showing the pixels used to define the spatial bins (see Section 4.1.1). While this figure shows the image-plane location, their relative location within the source plane of MRG-S0851 is noted in the legend.

populations. Through our spatially resolved analysis, we were not able to robustly constrain the ages of the two outer spatially resolved bins at  $<-1.4$  kpc and  $>1.5$  kpc, as our two choices of dust/metallicity SSP priors (see the two last columns of Table 1) lead to different ages. This is likely the result of the significant contamination in these outer grism spectra by the cluster member galaxies and other nearby objects, yielding an average  $S_{\text{grism}}/N_{\text{contam}}$  ratio of  $\sim 0.2$  and  $1.1$  per pixel for these bins, where  $S_{\text{grism}}$  is the flux of MRG-S0851 and  $N_{\text{contam}}$  is the flux of other nearby objects in each pixel. For reference, the central bin has  $S_{\text{grism}}/N_{\text{contam}} \sim 22.8$  per pixel. Our analysis suggests that MRG-S0851 has a flat age gradient within the inner 3 kpc core, with a circularized effective radius of  $1.7^{+0.3}_{-0.1}$  kpc in the source plane (see Appendix B, for size measurements); it is similarly compact relative to the population of quiescent galaxies with the same stellar mass and redshift (e.g., van der Wel et al. 2014). With a median age of 1.8 Gyr, it seems to be older than similar quiescent galaxies selected from the Illustris simulation, which have  $t_{50} = 1.1^{+0.3}_{-0.2}$  Gyr and  $\log M_*/M_\odot = 11.03^{+0.30}_{-0.37}$  (see Figure 7). MRG-S0851 is therefore consistent with an early formation scenario (e.g., Williams et al. 2014; Wellons et al. 2015). The recovered SFH gradients of MRG-S0851 will be discussed in further depth in a follow-up paper (M. Akhshik et al. 2020, in preparation).

We estimate the average mass-weighted age of  $1.9^{+0.2}_{-0.1}$  Gyr from our joint fit with `requiem2d`, consistent within the  $1\sigma$  uncertainty interval with the estimate of the average mass-weighted age from `Prospector- $\alpha$`  ( $2.0 \pm 0.3$  Gyr; Figure 11) and the median mass-weighted age obtained from the joint fit ( $1.8^{+0.3}_{-0.2}$  Gyr). We estimate a stellar mass of  $\log M_*/M_\odot = 11.02 \pm 0.04$  from the joint fit, consistent with the `Prospector- $\alpha$`  estimate within  $1\sigma$  (Section 6.1). Finally, the sSFR is estimated to be  $\log \text{sSFR}_{100 \text{ Myr}}/[\text{yr}^{-1}] = -10.32^{+0.07}_{-0.05}$  from the spectrophotometric fit. It is  $\sim 0.5$  dex lower than the sSFR estimate of

`Prospector- $\alpha$`  from photometry only. We discuss this discrepancy in Section 7.

In the G141 bandpass at the redshift of MRG-S0851  $z = 1.88$ , we observe a few spectral features that are sensitive to the age of stellar populations. Most notably, we sample the 4000 Å break, Mg I and the Balmer lines  $H\beta$ ,  $H\gamma$ , and  $H\delta$ . As the Balmer lines appear to be filled by emission, they are not expected to drive the age fit significantly. We detect [O III] emission lines at the rest-frame wavelengths of 5008.2 Å and 4960.3 Å (to be discussed further in Section 6.3).

As discussed in Section 4.3, we add the emission line templates using `Grizli` together with the SSPs that do not include nebular emission lines that are generated by FSPS. Another option is to instead use FSPS to generate nebular emission lines. To test this alternative approach, we use the `Prospector- $\alpha$`  posterior of gas-phase metallicity along with the stellar metallicity and dust parameters to add nebular emission lines, specifically adding more templates between 1 and 10 Myr in logarithmic lookback time steps. We confirm that the resulting global and resolved ages are all consistent in both approaches within  $1\sigma$  statistical uncertainty—except for bins at  $\leq -1.4$  kpc and  $\geq 1.5$  kpc, due to contamination. The global mass-weighted median age using FSPS templates that include nebular emission lines is constrained to be  $1.9 \pm 0.2$  Gyr.

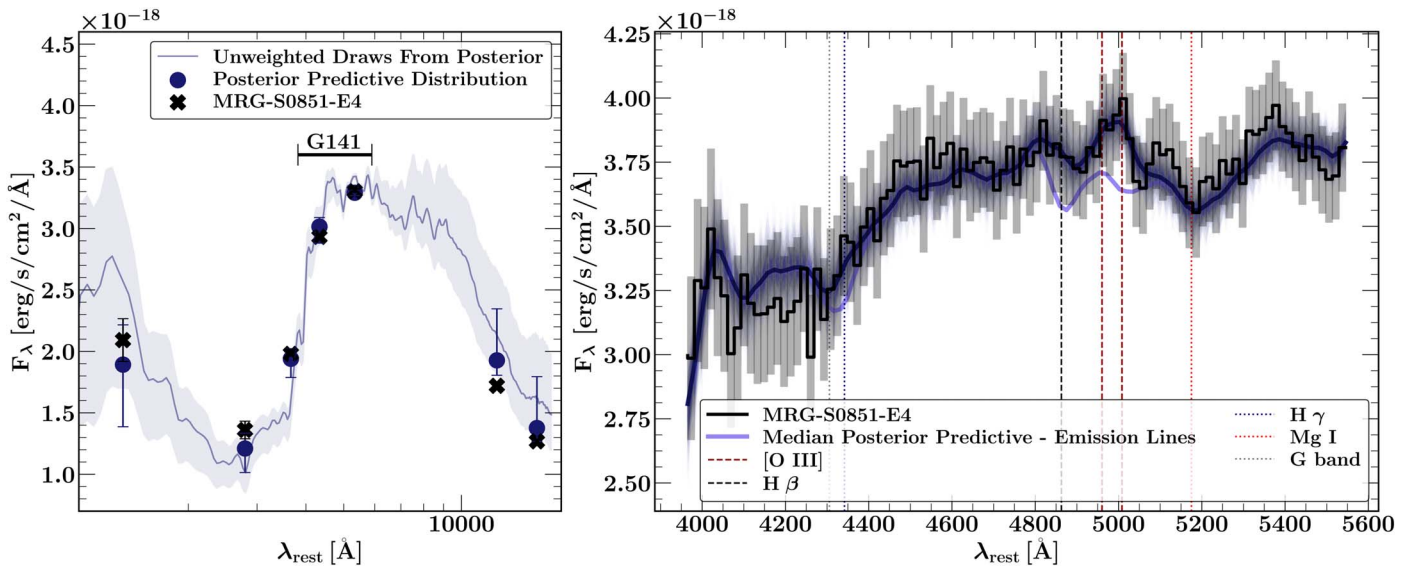
When accounting for the effects of dust and metallicity on the SSPs of resolved stellar populations of MRG-S0851, the distribution of resolved ages may change for different priors (Figure 15, comparing black and red points; also see Section 4.2.2 for further discussion of these priors). We cannot calculate the resolved photometry for the two Spitzer IRAC channels, and as a result, neither the resolved joint fit nor the resolved `Prospector- $\alpha$`  fit fully includes resolved rest-frame near-IR photometry.<sup>23</sup> Therefore, the priors on SSPs, which control dust and metallicity, may play a significant role and affect the weights  $x$  of SSPs. While this effect is not large for the five central bins, it seems to be more significant for the outer bins at  $\leq -1.4$  kpc and  $\geq 1.5$  kpc, where the SSP priors are strongly driving the fit and yielding different results for the ages of these bins, as the photometry error bars are larger and grism spectroscopy is significantly contaminated.

### 6.3. Emission-line Diagnostics of MRG-S0851

We fit for the fluxes of emission lines, as described in Sections 4.3 and 6.2. The statistical significance of the spatially unresolved emission lines is estimated to be  $\sim 9\sigma$  ([O III]),  $\sim 5\sigma$  ( $H\beta$ ),  $\sim 3\sigma$  ( $H\gamma$ ), and  $\sim 1\sigma$  ( $H\delta$ ) from the corresponding Markov Chains, and we therefore detect [O III],  $H\beta$ , and  $H\gamma$  with  $\gtrsim 3\sigma$  statistical certainty.

[O III] is detected in three central bins ( $>3\sigma$ ). We constrain the global  $\log([\text{O III}]/H\beta)$  to be  $0.24^{+0.10}_{-0.08}$ . Taken together, it is likely that the emission lines originate from star formation activity. However, it may be that MRG-S0851 has a high [N II]/ $H\alpha$  ratio; combined with low [O III]/ $H\beta$ , this would favor an AGN (Baldwin et al. 1981) or a low-ionization emission-line region (e.g., Heckman 1980; Belfiore et al. 2016). Unfortunately, both [N II] and  $H\alpha$  are outside of the

<sup>23</sup> As a reminder, in our spatially resolved joint fit, we require that the sum of the predicted resolved IRAC photometric fluxes in our model match the observed global values (Section 4.1.2).



**Figure 13.** Left: global spectral energy distribution (SED) of MRG-S0851. Black crosses are the HST and Spitzer photometric data. Blue circles are the posterior predictive distribution. Dark blue curve and its light blue shade are the unweighted median and 16th–84th percentiles of the draws from posterior (i.e., unlike the blue circles, they are not weighted by the weights of the Dirichlet process). Right: global extracted 1D grism spectra of MRG-S0851, following Horne (1986) for the 1D extraction. Various spectral features in the G141 bandpass at the redshift of MRG-S0851 are shown by vertical lines. Data are shown in black, with gray vertical bars indicating the  $1\sigma$  errors. Dark blue curves are draws from the posterior predictive distribution. We smooth both the data and models using the `Gaussian_filter` method of `scipy` with  $\sigma = 0.3$ . We also plot the full median model with a darker blue line and the median continuum-only model removing emission lines with a lighter blue line.

WFC3/G141 bandpass (and unresolvable) for MRG-S0851 to examine this scenario.

In the spatially resolved analysis, we also fit for the flux of [O III],  $H\beta$ ,  $H\gamma$ , and even  $H\delta$  emission lines in the WFC3/G141 bandpass, but the weakness of these emission lines combined with the low spectral resolution of WFC3/G141 do not allow us to constrain any potential gradients in [O III]/ $H\beta$ , or statistically significant Balmer decrements.

## 7. Summary and Discussion

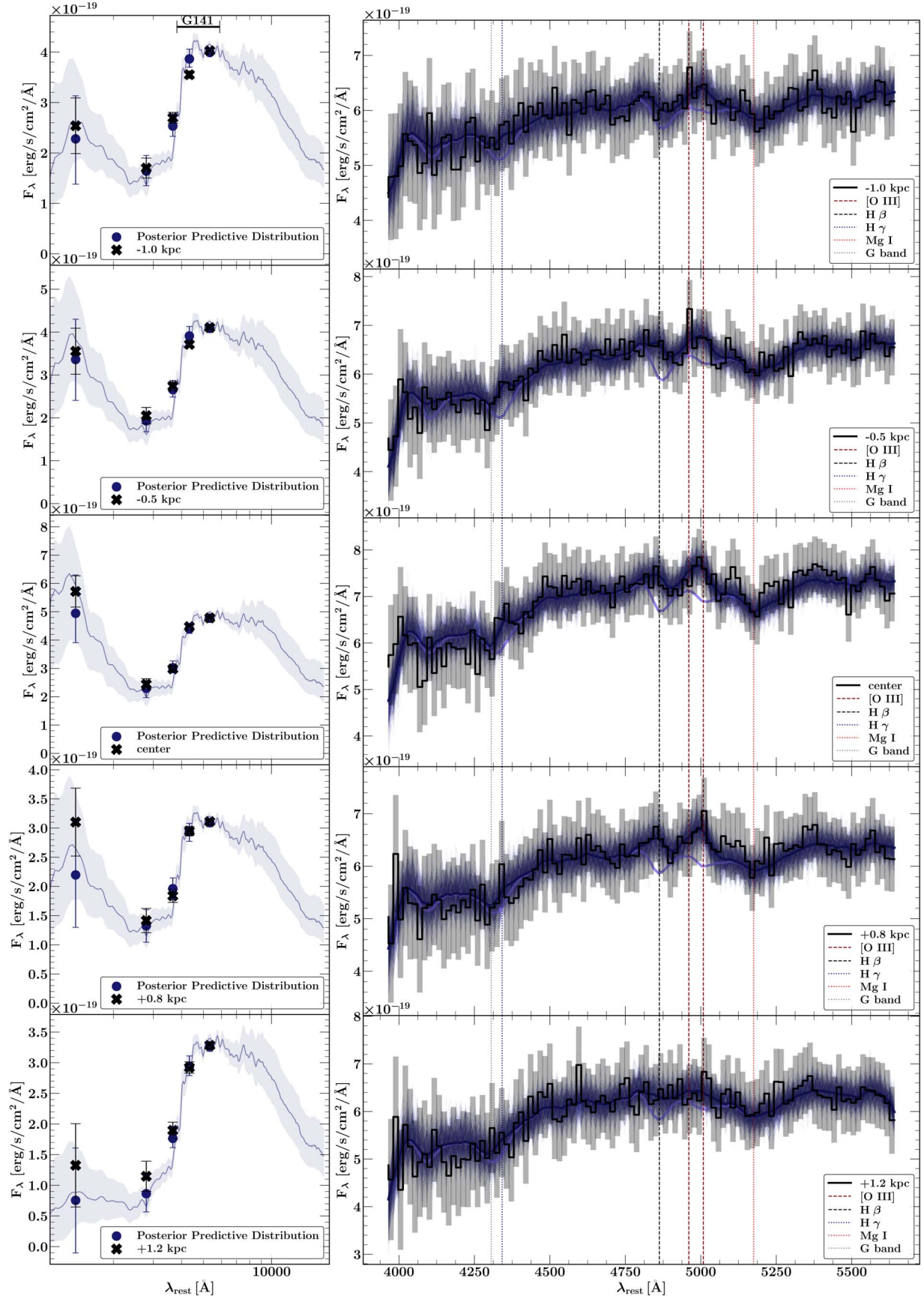
We present a methodology to jointly fit HST grism spectroscopy with HST and Spitzer photometry with the new `requiem2d` package, in order to constrain the age and SFH of stellar populations of galaxies. Our fitting method includes two steps, a preliminary fit with `Prospector- $\alpha$`  to photometric data, and a subsequent joint spectrophotometric fit using a linear combination of SSPs, generated by drawing from the `Prospector- $\alpha$`  posterior to constrain ages and SFHs. Our presentation here is tuned to HST grism spectroscopy and HST and Spitzer photometry; however, the core statistical model (Table 2 and Figure 4) can be applied to all well-calibrated spectroscopic and photometric data. The beta version release of the `requiem2d` code accompanies this manuscript, accessible through a public GitHub repository.<sup>24</sup>

We test our methodology using a sample of massive  $\log M_*/M_\odot \geq 10.6$  quiescent galaxies at  $z = 2$ , selected from the Illustris simulation, showing that the median mass-weighted age as well as the general trends of SFH can be recovered with no noticeable biases. We also use this method in a pilot study to analyze the spatially resolved stellar populations of a lensed massive red galaxy, MRG-S0851, using photometric data taken using HST/WFC3 broadband filters covering the rest-

frame UV to optical, HST/WFC3 G141 grism data, and Spitzer IRAC channels 1 and 2 data. With grism spectroscopy, we constrain the redshift to be  $z = 1.883 \pm 0.001$  by fitting MRG-S0851 using `Grizli`. By constructing a consistent lensing model, we correct the stellar mass for the effects of strong gravitational lensing, constraining the stellar mass to be  $\log M_*/M_\odot = 11.02 \pm 0.04$  from our joint fit with `requiem2d`. The circularized effective radius is measured to be  $r_c = 1.7^{+0.3}_{-0.1}$  kpc ( $0''.21 \pm 0''.02$ ) in the source plane (see Appendix B.1). We fit the global dust and metallicity using `Prospector- $\alpha$` , with the results reported in Table 3. From a joint spectrophotometric analysis, we find that the unresolved stellar populations have a global median mass-weighted age of  $1.8^{+0.3}_{-0.2}$  Gyr, a global light-weighted age of  $1.4^{+0.2}_{-0.1}$  Gyr, and a global sSFR of  $\log \text{sSFR}_{100 \text{ Myr}}/[\text{yr}^{-1}] = -10.32^{+0.07}_{-0.05}$ . The sSFR obtained using `requiem2d` from the joint fit is 0.5 dex lower than the `Prospector- $\alpha$`  sSFR from a fit to photometry alone (Figure 11). We confirm that the difference is a result of adding grism spectroscopy, and fitting photometry alone using `requiem2d` yields results consistent with those of `Prospector- $\alpha$` . In fact, by adding grism spectroscopy that lacks the relatively strong Balmer absorption lines of ages  $\sim 0.1$ –1 Gyr, the SFR in this range of lookback time decreases while the SFR at  $\geq 1$ –2 Gyr increases, leading to an overall lower sSFR in the last  $\sim 100$  Myr of evolution. The change in SFH in the joint spectrophotometric versus photometry-only fit therefore mostly affects lookback-time ranges of 0.1–2 Gyr. However, this change in the overall shape of the inferred SFH does not significantly change the median/average mass-weighted ages of MRG-S0851, as these ages are dominated by the older SSP templates with lookback times of  $\gtrsim 2$  Gyr.

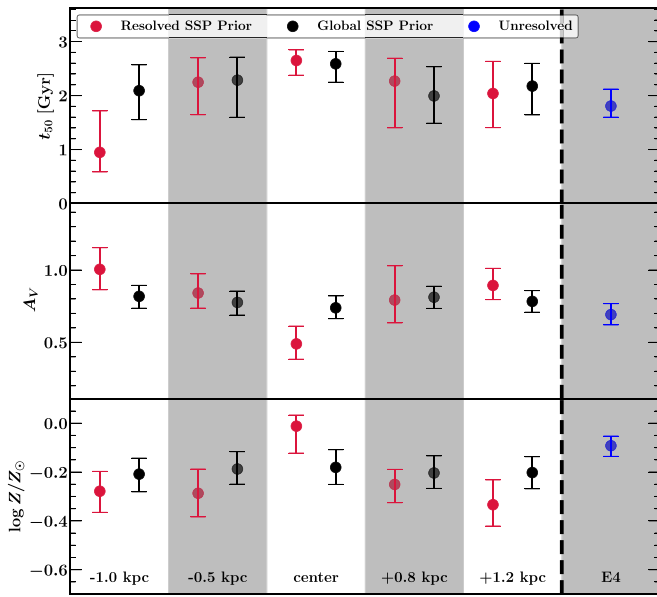
Leveraging the strong gravitational lensing magnification, we define seven spatial bins in order to study the spatially resolved stellar populations and to measure the age gradient

<sup>24</sup> <https://github.com/makhshik/requiem2d>



**Figure 14.** Resolved extracted 1D grism spectra and the SED of MRG-S0851, obtained using the global SSP prior for dust and metallicity (Section 4.2). We smooth both the data and the posterior predictive distribution using the `Gaussian_filter` method of `scipy` with  $\sigma = 0.3$ . Symbols and 1D extraction method are the same as in Figure 13.





**Figure 15.** Resolved median ages,  $A_V$ , and metallicities of central spatial bins (left) as well as the unresolved global values MRG-S0851 (right). For the spatial bins, we plot two sets of mass-weighted average ages obtained using two different dust/metallicity priors (Section 4.2). One set is obtained by using the resolved *Prospector- $\alpha$*  dust/metallicity posteriors as priors to generate SSPs in the joint fitting (dubbed Resolved SSP Prior), and the other is obtained by using the global *Prospector- $\alpha$*  dust/metallicity posterior as a prior, generating SSPs for all spatially resolved bins (dubbed Global SSP Prior). Blue circles in right panel are the global values of MRG-S0851, as measured in the unresolved analysis.

(Figure 15). We defer the discussion of the SFH gradient to a subsequent paper. At face value, the flat age gradient combined with a relatively old global age and compact size favors an early formation scenario for MRG-S0851 (e.g., Williams et al. 2014; Wellons et al. 2015).

Emission-line diagnostics in the HST/WFC3 G141 band-pass at the redshift of MRG-S0851 disfavor the presence of an AGN at the center of the galaxy, as the [O III] emission line is not centrally concentrated and the ratio of [O III]/H $\beta$  is not high. While the [O III]/H $\beta$  ratio is consistent with star formation activity, we caution that we need to observe H $\alpha$  and [N II] to rule out AGN or low-ionization emission-line regions. These emission lines can be studied easily in the near future with the NIRSpec/Integrated Field Unit of the James Webb Space Telescope (JWST). These lines are also accessible to ground-based spectroscopy using Keck/MOSFIRE, albeit at a lower resolution than is provided by JWST. Observations of H $\alpha$  will also be helpful to constrain the instantaneous SFR.

MRG-S0851 is the first target in the ongoing REQUIEM galaxy survey, which includes grism spectroscopy using HST/WFC3 G141 as a part of the HST-GO-15663 ongoing program, for a sample of eight strongly lensed quiescent galaxies spanning a redshift range of  $1.6 < z < 2.9$  and a stellar mass range of  $10.4 < \log M_*/M_\odot < 11.7$  (HST-GO-15633). The analysis of the spatially resolved stellar populations of the rest of the REQUIEM targets will follow the framework developed herein.

We would like to thank the anonymous referee for sharing valuable comments and questions that helped us to significantly improve the draft. We gratefully acknowledge support by NASA under award No 80NSSC19K1418, HST-GO-14622, and

HST-GO-15663. This work is based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Financial support for M.A. and K.W. is gratefully acknowledged. K.W. wishes to acknowledge funding from the Alfred P. Sloan Foundation. S.T. acknowledges support from the ERC Consolidator Grant funding scheme (project ConText, grant No. 648179). The Cosmic Dawn Center is funded by the Danish National Research Foundation under grant No. 140. S.T. and G.B. acknowledge support from the ERC Consolidator Grant funding scheme (project ConText, grant No. 648179). C.C.W. acknowledges support from the National Science Foundation Astronomy and Astrophysics Fellowship grant AST-1701546. The Dunlap Institute is funded through an endowment established by the David Dunlap family and the University of Toronto.

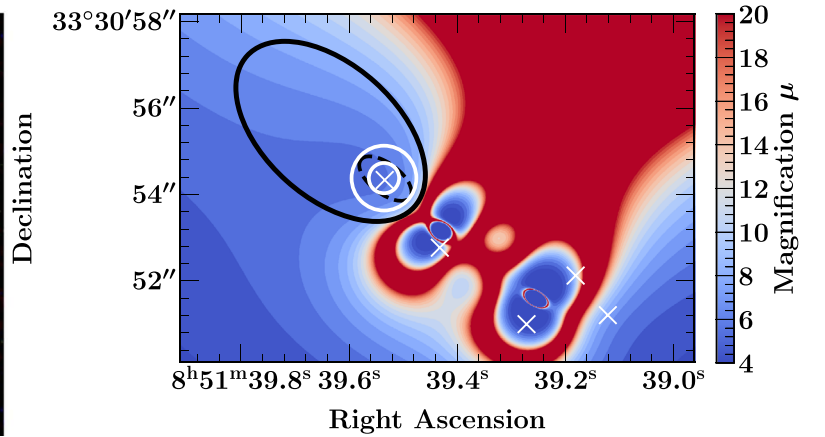
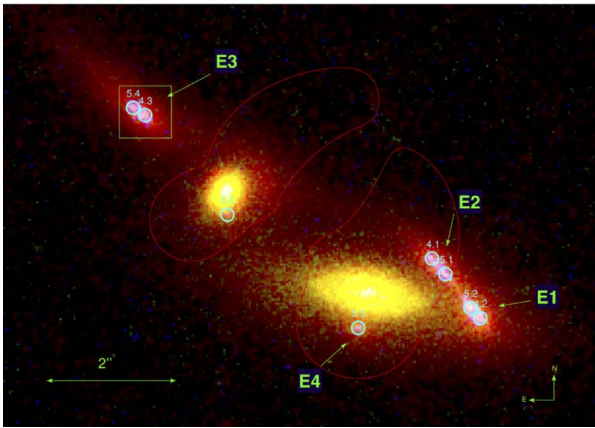
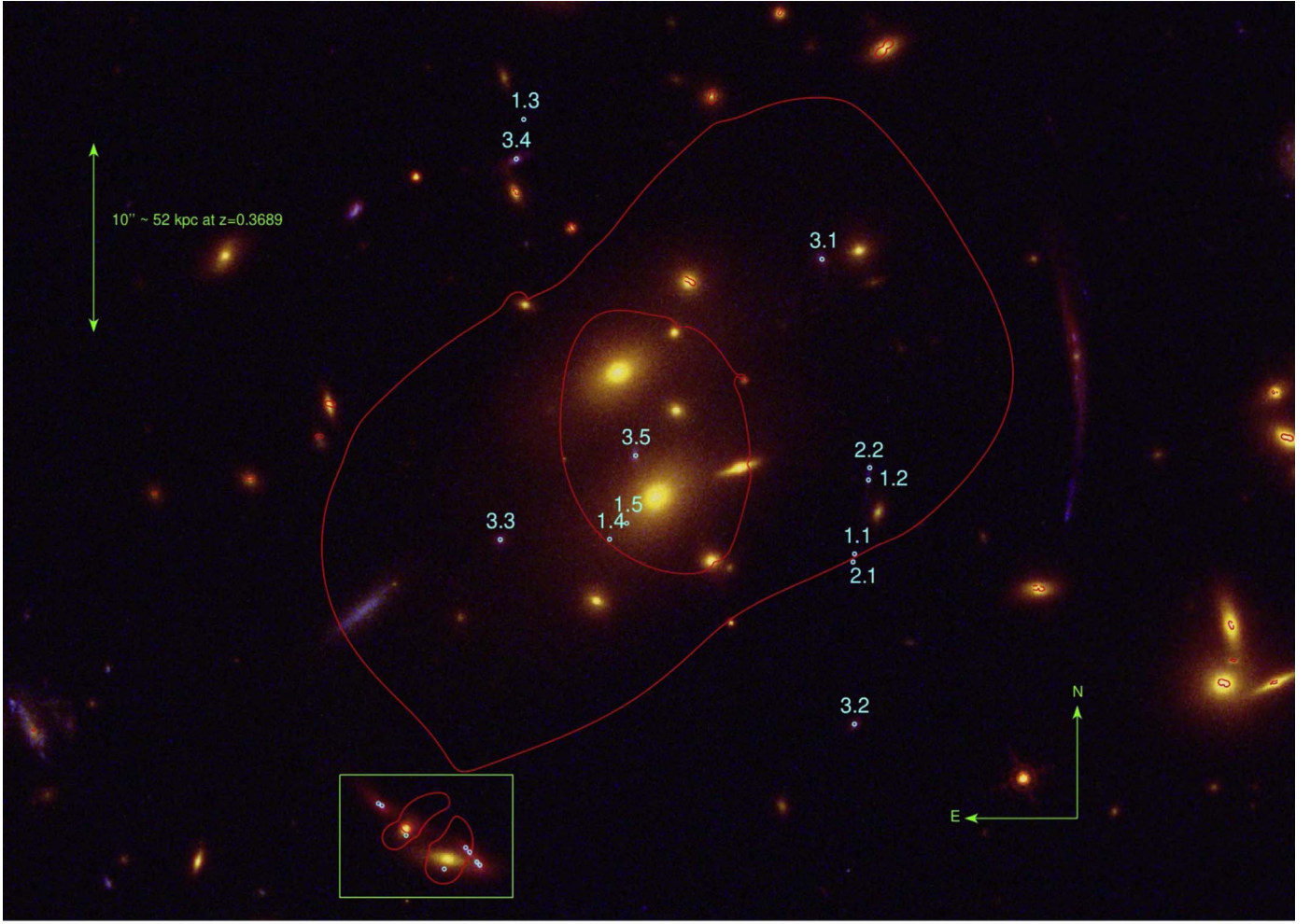
## Appendix A Gravitational Lens Modeling of MRG-S0851

We highlight the key aspects of the MRG-S0851 lens model. This lens model is used to transform information to the source plane and interpret the results of stellar population analyses.

### A.1. Methodology of Lensing Mass Models

While we provide a brief summary of the gravitational lensing analysis used in this work here, we refer the reader to Kneib & Soucaill (1996), Richard et al. (2010), Verdugo et al. (2011), and Smith et al. (2015) for a more in-depth discussion of the lensing algorithm used here. We adopt a parametric approach using *Lenstool* (Jullo et al. 2007) to model the cluster mass distribution surrounding our target as a combination of dual pseudo-isothermal ellipsoids (dPIEs; Elíasdóttir & Möller 2007), using a Monte Carlo Markov Chain method to estimate the parameters and their uncertainties. These dPIE clumps are combined to map the dark matter (DM) at the cluster scale and to model the cluster mass distribution, and galaxy-scale DM potentials are used to describe galaxy-scale substructure. Given the large number of galaxies in the cluster, it is not feasible to optimize the parameters of every potential, as the large parameter space will lead to an unconstrained minimization. Moreover, individual galaxies do not contribute significantly to the total mass budget of the cluster, and their effects on lensing are minimal unless they are projected close to the lensed galaxies. To reduce the overall parameter space, we therefore scale the parameters of each galaxy using a reference value with a constant mass–luminosity scaling relation (see Limousin et al. 2007).

We construct a galaxy cluster catalog using the red sequence technique (Gladders & Yee 2005), where we select galaxies that have similar colors in the  $J_{F814W} - J_{F125W}$  color versus  $J_{F125W}$ -band magnitude diagram. Our original catalog includes 136 cluster members. As the bright cluster galaxies (BCGs) of galaxy clusters do not follow the red sequence, we remove the two BCGs (Newman et al. 2013a, 2013b) from the galaxy catalog and model them separately. To allow for estimation of lensing magnification, we also decide to remove from the galaxy catalog the two cluster members responsible for the main perturbation of the lensed object and model them separately, too. Additionally, an intrinsic scatter, which is expected in the mass–luminosity relation, offers further



**Figure A1.** Top panel: composite color of SDSS J0851+3331, from HST/WFC3  $H_{F160W}$  (red),  $I_{F814W}$  (green), and  $U_{F390W}$  (blue). Red lines show the critical curves of the gravitational potential at the redshifts of the lensed object  $z = 1.88$ . Cyan circles mark the constraints used to model the system as described in Appendix A.1. Bottom left panel: zoom-in on a region where images of MRG-S0851 appear. The five images are labeled with E1-5. Green box around image E3 shows the region where the magnification is computed to be  $\mu = 5.7^{+0.4}_{-0.2}$ . Bottom right panel: the magnification map in the region containing the lensed images of MRG-S0851. Solid black ellipse outlines the segmentation map. Dashed black ellipse shows the half-light radius. Magnification is almost constant across image 3;  $\mu = 6.2^{+1.2}_{-0.7}$  in the solid black ellipse (segmentation map) and  $\mu = 5.7^{+0.7}_{-0.2}$  in the dashed black ellipse (half-light radius). White circles show the  $0.7''$  and  $1.5''$  apertures that we adopted to measure global photometry. We indicate the location of the different MRG-S0851 images with crosses.

physical motivations to individually model the galaxies by including enough constraints.

Our final lens model of MRG-S0851 includes two cluster-scale DM halos parameterized as dPIE profiles. For minimization

procedures, we let all the parameters of halo vary, with the exception of the truncation radius  $r_{\text{cut}}$ , because it extends beyond the strong lensing regime and therefore cannot be constrained. The final positions of the DM clumps remain close to their BCG

**Table A1**  
Position of the Constraints Used to Construct the Lensing Models

ID	R.A.	Decl.	Notation in Sharon et al. (2020)	$z$	$z$ Reference
1.1	8:51:38.0254	+33:31:03.132	B1.1	1.3454	Bayliss et al. (2011)
1.2	8:51:37.9667	+33:31:07.032	B1.2	1.3454	Bayliss et al. (2011)
1.3	8:51:39.4200	+33:31:26.034	B1.3	1.3454	Bayliss et al. (2011)
1.4	8:51:39.0580	+33:31:03.910	B1.4	1.3454	Bayliss et al. (2011)
1.5	8:51:38.9859	+33:31:04.751	B1.5	1.3454	Bayliss et al. (2011)
2.1	8:51:38.0315	+33:31:02.708	B2.1	1.3454	Bayliss et al. (2011)
2.2	8:51:37.9613	+33:31:07.675	B2.2	1.3454	Bayliss et al. (2011)
3.1	8:51:38.1625	+33:31:18.666	D3.1	1.79	This work
3.2	8:51:38.0261	+33:30:54.162	D3.2	1.79	This work
3.3	8:51:39.5185	+33:31:03.890	D3.3	1.79	This work
3.4	8:51:39.4513	+33:31:23.948	D3.4	1.79	This work
3.5	8:51:38.9480	+33:31:08.315	D3.5	1.79	This work
4.1	8:51:39.6640	+33:30:47.657	E2.1	1.88	This work
4.2	8:51:39.6045	+33:30:46.741	E2.2	1.88	This work
4.3	8:51:40.0173	+33:30:49.868	E2.3	1.88	This work
5.1	8:51:39.6474	+33:30:47.418	E1.2	1.88	This work
5.2	8:51:39.6168	+33:30:46.898	E1.1	1.88	This work
5.3	8:51:39.7549	+33:30:46.532	E1.4	1.88	This work
5.4	8:51:40.0323	+33:30:49.980	E1.3	1.88	This work
5.5	8:51:39.9154	+33:30:48.302	E1.5	1.88	This work

**Note.** The coordinates are reported in Sexagesimal coordinates (adopting J2000 epoch). We note that constraints 4.1–4.3 and 5.1–5.5 in this table are labeled as Figure A1. We measure the redshifts of each constraint in systems D and E individually, using the default redshift fitting methods of `Grizli`. We use 12 orbits of grism WFC3/G141 data, obtained as a part of the program HST-GO-14622 (PI: K. Whitaker), for fitting the redshifts of systems D and E. Two constraints 3.2 and 3.4 have clean grism spectra from the program, HST-GO-14622, and the redshift fitting with the grism data was inconclusive for the rest of the constraints 3.1, 3.3, and 3.5 due to contamination of their spectra by nearby objects. The redshift fit of constraints 3.2 and 3.4 is driven both by [O III] emission-line doublet at a rest-frame wavelength of 5008 Å, yielding a reduced  $\chi^2$  of 0.97 and 0.94, respectively.

( $\leq 5''$ ). We also tested a model with only one cluster-scale DM halo, but the result was not as good as that obtained with two cluster-scale DM halos (see Appendix A.2).

We constrain the cluster using the three-lensed systems that have confirmed spectroscopic redshifts from Sharon et al. (2020) (Figure A1, left panel). In the two of lensed systems, we identify resolved emission knots within their image and use them as additional constraints. Figure A1 shows the position of the constraints (see Table A1 for the exact coordinates).

#### A.2. Modeling Results, Choice of the Best Model

To estimate a reliable magnification for MRG-S0851, we try several different models: we change the number of cluster-scale DM halos, and we also add a fifth image, denoted as system 5.5 (see Table A2). We quantitatively compare the quality of different models with two criteria. The first criterion is the root mean square (rms), which describes how well the model reproduces the positions of the constraints. The second criterion is the Bayesian Information Criterion (BIC), which is a statistical measurement based on the model likelihood, penalized by the number of free parameters and the number of constraints (see, e.g., Limousin et al. 2010; Mahler et al. 2018). We list the results in Table A2. Models with two DM halos perform significantly better than the models with one DM halo, based on the rms criterion. This effect can be explained by the additional flexibility. However, as the BIC does not change significantly, we conclude that it is balanced by the increase in goodness of the fit for all models with two DM halos. The only exception is the two DM halo model with the fifth image as a part of system 4, and we therefore reject it. Finally, we reject the model that shows the best rms (the

**Table A2**  
Summary of the Different Modeling Assumptions and the Criteria Used for Comparison

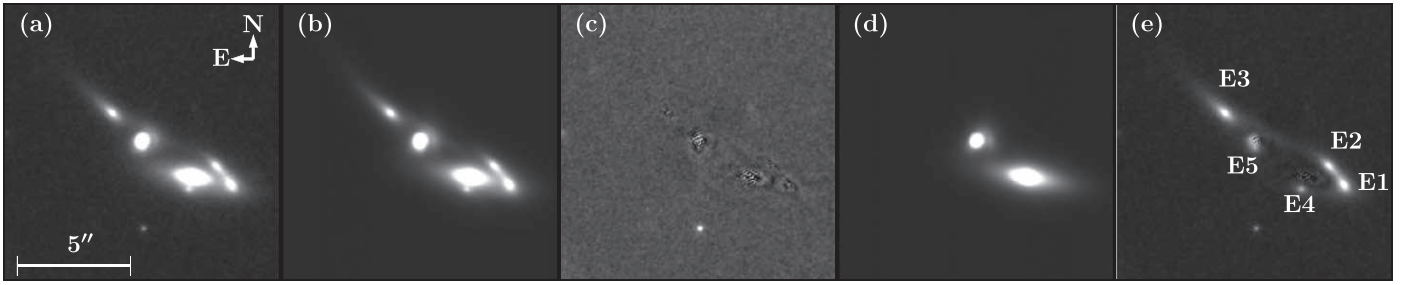
Model Number	Cluster Scale DM Halo	Fifth Image Assumption	Rms	BIC	$\mu$
1	2 DM	Part of system 5	0''21	87	$5.72^{+0.36}_{-0.2}$
2	2 DM	Not used	0''19	84	$5.96^{+0.22}_{-0.19}$
3	2 DM	Part of system 4	0''29	96	$8.45^{+0.9}_{-0.57}$
4	1 DM	Part of system 5	0''37	84	$8.38^{+0.62}_{-0.46}$
5	1 DM	Not used	0''32	77	$6.45^{+0.22}_{-0.19}$
6	1 DM	Part of system 4	0''36	84	$5.12^{+0.23}_{-0.19}$

**Note.** The rms describes how well the model reproduces the positions of the constraints. The Bayesian Information Criterion (BIC) is a statistical measurement based on the model likelihood, penalized by the number of free parameters and the number of constraints (see, e.g., Limousin et al. 2010). The last column denotes the estimated magnification,  $\mu$ , for each model within the green box in the bottom left panel of Figure A1.

lowest) and best BIC (the lowest), because the mass of the nearby cluster member galaxy (the one located to the east in Figure A1, bottom left panel) is unnaturally low. We therefore keep the second-best model, with two DM halos and the fifth image as part of system 5, as the best physical model.

Using our final model, we compute the average gravitational magnification in a representative area of the galaxy to be  $\mu = 5.7^{+0.4}_{-0.2}$  within the green box in the bottom left panel of Figure A1. The uncertainty denotes the  $1\sigma$  width of the distribution of  $\mu$  for all pixels in the box. We also calculate the magnification in the same box for all models in Table A2.





**Figure A2.** (a) Cutout image of MRG-S0851 with the nearby subcluster members. (b) GALFIT model of all light profiles in the field. (c) Residual of the GALFIT model of all light profiles. (d) GALFIT model of the subcluster members. (e) Quintuple images of MRG-S0851 after removing subcluster lenses using their GALFIT model. The five multiple images of MRG-S0851 are labeled in panel (e).

As can be seen in Table A2, different models lead to different median values of the magnification, and the scatter of the median values for different models is greater than the statistical uncertainty of the best model. We also calculated the light-weighted magnification, the average value of the magnification for all pixels of E3 weighted by its light profile, and we find a mean and a standard deviation of  $\mu_l = 7.6 \pm 2.1$  for all models. However, as we discussed earlier, these models can be ruled out based on their physical prediction for the mass of nearby cluster members and/or a combination of the rms and BIC. We therefore suggest that the scatter in the median value of all models and the light-weighted magnification is an overestimate, and we use the  $1\sigma$  width of the distribution of  $\mu$  for all pixels in the representative area as an estimate of uncertainty in the rest of the paper.

Sharon et al. (2020) also present a lensing model for the cluster SDSS J0851+3331. We note that, while Sharon et al. (2020) optimize the global properties of the cluster and source positions, we focus on reproducing the geometry of one specific system E, our main science target. In this work, we construct six different models (summarized in Table A2) to study and understand this specific system better. Specifically, the main difference between the models presented herein and Sharon et al. (2020) is the added flexibility in the vicinity of the lensed system E. Here, we allow all of the parameters of the galaxies near the systems to vary except for their  $x$  and  $y$  positions, while Sharon et al. (2020) fix  $x$ ,  $y$ ,  $e$ , and  $\theta$  to their observed properties. Sharon et al. (2020) also model the core of the cluster with one dominant cluster-scale halo, whereas we find statistical evidence that modeling the core with two halos is favorable.

## Appendix B MRG-S0851 Morphology

In this section, we present our method to constrain the light profile of MRG-S0851 and measure its effective radius. We focus our analyses on E3, the third image (see Figure 1), because it is the brightest image ( $m_{H_{F160W}} = 20.36$ ) and the grism spectrum is less contaminated by nearby objects.

### B.1. The Image-plane GALFIT Model

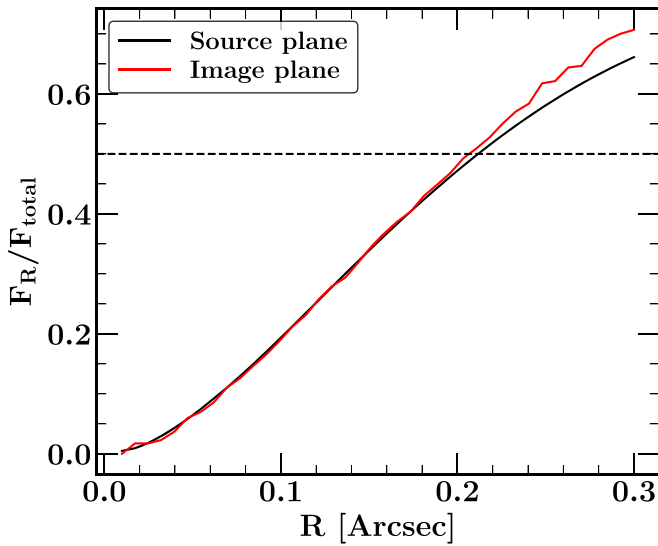
To study the size and the morphology of the different images of MRG-S0851, we use the PSFs and the weight maps that are generated as parts of the data reduction and the photometric analysis together with the  $H_{F160W}$  drizzled mosaic as basic inputs in the GALFIT software (Peng et al. 2011). MRG-S0851 is close to subcluster members on the sky, with a significant overlapping of the light profiles. We therefore model out all

light profiles in a  $12'' \times 12''$  box around the center of the subcluster members, using an iterative method to improve our model. The Source Extractor catalog coordinates are used as initial values for the center of the light profiles, the initial Sérsic index value is  $n = 4$ , and we use the Source Extractor catalog magnitudes as our initial guess. We run GALFIT, inspect the residuals by eye, and add fainter components in the location with the highest residuals. We run GALFIT including the extra components, keeping the components only if there is an improvement in the reduced  $\chi^2$ . We fit 14 components in total, fixing the Sérsic index of two components as well as the axis ratio of a single component in our final iteration to avoid numerical instabilities. None of these components constitute the GALFIT model of E3, and as these components are  $\sim 1$  ABmag fainter than the GALFIT components of E3, it is unlikely that fixing these parameters affects the GALFIT model of E3 noticeably. The final result of the image-plane analysis is shown in Figure A2. We use the same procedure for the other four HST images, obtaining the GALFIT model for all five HST filters.

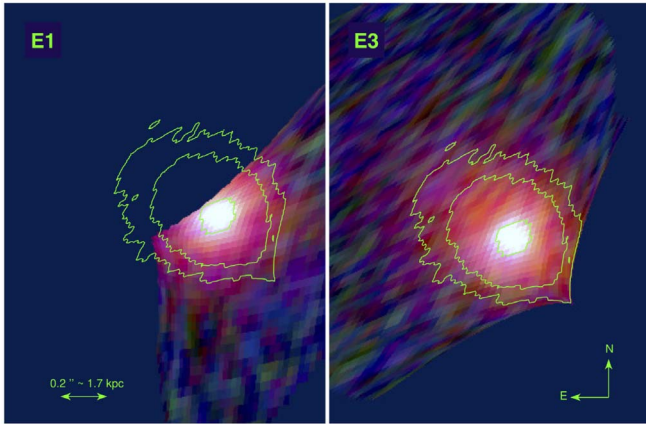
The GALFIT model of E3 includes two components with Sérsic indices of  $n_1 = 2.35 \pm 0.05$  and  $n_2 = 1.32 \pm 0.06$ ,  $H_{F160W}$  magnitudes of  $m_1 = 21.20 \pm 0.03$  and  $m_2 = 21.41 \pm 0.03$ , the image-plane semimajor axes of  $r_1 = 0''.43 \pm 0''.01$  and  $r_2 = 2''.72 \pm 0''.15$ , and the axis ratios of  $ar_1 = 0.299 \pm 0.003$  and  $ar_2 = 0.150 \pm 0.004$  (uncertainties are reported directly from GALFIT). One of these components corresponds to the extended tail of E3. These components are almost aligned with  $\delta\theta \sim 9^\circ$ . We estimate the spatial offset of 1.3 kpc in the source plane between these two components. The second extended, offset component is not required by the GALFIT model for any of the other HST filters, aside from  $H_{F160W}$ . We therefore suspect that this offset is a modeling artifact and not physical, caused by nontrivial gravitational lensing effects.

### B.2. The Source-plane Effective Radius

We next measure the effective radius in the source plane using the lensing model (see Appendix A for details of the lensing model). To isolate the effect of PSF in our source-plane measurements, we generate a PSF-deconvolved model for E3 and add the residuals, i.e., we calculate  $\text{Data} - \text{Model}_{\text{PSF-convolved}} + \text{Model}_{\text{PSF-deconvolved}}$ ; see Szomoru et al. (2012) and Newman et al. (2018a) for further discussions of this method. After mapping the resulting image back to the source plane, we define a set of circular apertures in the source plane with increasing radii and calculate the corresponding elliptical apertures in the image plane. We use these two equivalent



**Figure B1.** Ratio of the enclosed flux to the total flux of PSF-deconvolved images, measured using circular apertures in the source plane and the equivalent elliptical apertures in the image plane. Both ratios are reasonably consistent for different radii of interest, yielding the same half-light radius in both source- and image-plane measurements.



**Figure B2.** Source-plane reconstruction using the GALFIT model presented in Appendix B.1. We draw contours from image 3 and overplot them on image 1 after a simple shift. This figure shows that E1 is a partial image.

sets to measure the enclosed flux of the PSF-deconvolved images in both the source and image planes, obtaining a consistent result for the half-light radius, as can be seen in Figure B1. To measure the uncertainty of the half-light radius, we bootstrap 400 times and remeasure the half-light radius, and in each iteration, we randomize the  $H_{F160W}$ -band pixels using the HST weight-map uncertainty. The circularized effective radius of E3 is constrained to be  $r_c = 1.7^{+0.3}_{-0.1}$  kpc ( $0''.21 \pm 0''.02$ ). Our GALFIT model also shows that E1 and E2 are partially lensed images, so their light profiles are not representative of the whole system and can henceforth be ignored (Figure B2).

#### ORCID iDs

Katherine E. Whitaker <https://orcid.org/0000-0001-7160-3632>

Gabriel Brammer <https://orcid.org/0000-0003-2680-005X>

Guillaume Mahler <https://orcid.org/0000-0003-3266-2001>

Keren Sharon <https://orcid.org/0000-0002-7559-0864>  
 Joel Leja <https://orcid.org/0000-0001-6755-1315>  
 Matthew B. Bayliss <https://orcid.org/0000-0003-1074-4807>  
 Rachel Bezanson <https://orcid.org/0000-0001-5063-8254>  
 Erica J. Nelson <https://orcid.org/0000-0002-7524-374X>  
 Jane R. Rigby <https://orcid.org/0000-0002-7627-6551>  
 Sune Toft <https://orcid.org/0000-0003-3631-7176>  
 Christina C. Williams <https://orcid.org/0000-0003-2919-7495>

#### References

- Abramson, L. E., Newman, A. B., Treu, T., et al. 2018, *AJ*, 156, 29  
 Allam, S. S., Tucker, D. L., Lin, H., et al. 2007, *ApJL*, 662, L51  
 Avila, R. J., & Hack, W. J. 2012, AAS Meeting, 220, 135.13  
 Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5  
 Bayliss, M. B., Hennawi, J. F., Gladders, M. D., et al. 2011, *ApJS*, 193, 8  
 Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, *MNRAS*, 461, 3111  
 Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, 874, 17  
 Belokurov, V., Evans, N. W., Hewett, P. C., et al. 2009, *MNRAS*, 392, 104  
 Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393  
 Brammer, G. 2016, hst prop, 14553  
 Brammer, G. B., Sánchez-Janssen, R., Labbé, I., et al. 2012a, *ApJL*, 758, L17  
 Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503  
 Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012b, *ApJS*, 200, 13  
 Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, *ApJ*, 739, 24  
 Casertano, S., de Mello, D., Dickinson, M., et al. 2000, *AJ*, 120, 2747  
 Chabrier, G. 2003, *PASP*, 115, 763  
 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102  
 Connor, R. J., & Mosimann, J. E. 1969, *J. Am. Stat. Assoc.*, 64, 194  
 Conroy, C., & Gunn, J. E. 2010, FSPS: Flexible Stellar Population Synthesis v3.1, Astrophysics Source Code Library, ascl:1010.043  
 Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486  
 Cook, B. A., Conroy, C., van Dokkum, P., & Speagle, J. S. 2019, *ApJ*, 876, 78  
 Creutz, M. 1988, *PhRvD*, 38, 1228  
 D'Eugenio, C., Daddi, E., Gobat, R., et al. 2020, *ApJL*, 892, L2  
 Dressler, A., Kelson, D. D., & Abramson, L. E. 2018, *ApJ*, 869, 152  
 Ebeling, H., Stockmann, M., Richard, J., et al. 2018, *ApJL*, 852, L7  
 Elíasdóttir, Á., & Möller, O. 2007, *JCAP*, 2007, 006  
 Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2019, *ApJ*, 870, 133  
 Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2020, *ApJ*, 898, 171  
 Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2016, arXiv:1612.05243  
 Foreman-Mackey, D., Sick, J., & Johnson, B. 2014, python-fsps: Python bindings to FSPS v0.1.1, Zenodo, doi:10.5281/zenodo.12157  
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1  
 Geier, S., Richard, J., Man, A. W. S., et al. 2013, *ApJ*, 777, 87  
 Gelman, A., & Rubin, D. B. 1992, *StaSci*, 7, 457  
 Genel, S., Vogelsberger, M., Springel, V., et al. 2014, *MNRAS*, 445, 175  
 Gladders, M. D., Rigby, J. R., Sharon, K., et al. 2013, *ApJ*, 764, 177  
 Gladders, M. D., & Yee, H. K. C. 2005, *ApJS*, 157, 1  
 Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, *Natur*, 428, 625  
 Heckman, T. M. 1980, *A&A*, 500, 187  
 Hennawi, J. F., Gladders, M. D., Oguri, M., et al. 2008, *AJ*, 135, 664  
 Hill, A. R., Muzzin, A., Franx, M., & van de Sande, J. 2016, *ApJ*, 819, 74  
 Homan, M. D., & Gelman, A. 2014, *J. Mach. Learn. Res.*, 15, 1593  
 Home, K. 1986, *PASP*, 98, 609  
 Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2012, *ApJ*, 752, 95  
 Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2015, *ApJ*, 802, 127  
 Jones, T., Ellis, R., Jullo, E., & Richard, J. 2010, *ApJL*, 725, L176  
 Jullo, E., Kneib, J. P., Limousin, M., et al. 2007, *NJPh*, 9, 447  
 Kelson, D. D., Williams, R. J., Dressler, A., et al. 2014, *ApJ*, 783, 110  
 Kneib, J. P., & Soucail, G. 1996, in IAU Symp. 173, Astrophysical Applications of Gravitational Lensing, ed. C. S. Kochanek & J. N. Hewitt (Dordrecht: Kluwer), 113  
 Koester, B. P., Gladders, M. D., Hennawi, J. F., et al. 2010, *ApJL*, 723, L73  
 Kriek, M., & Conroy, C. 2013, *ApJL*, 775, L16  
 Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJL*, 649, L71  
 Kron, R. G. 1980, *ApJS*, 43, 305  
 Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, *ApJL*, 649, L67  
 Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. 2016, *ApJ*, 831, 152  
 Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019a, *ApJ*, 876, 3

- Leja, J., Johnson, B. D., Conroy, C., et al. 2019b, *ApJ*, **877**, 140
- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, **837**, 170
- Limousin, M., Jullo, E., Richard, J., et al. 2010, *A&A*, **524**, A95
- Limousin, M., Richard, J., Jullo, E., et al. 2007, *ApJ*, **668**, 643
- Lin, H., Buckley-Geer, E., Allam, S. S., et al. 2009, *ApJ*, **699**, 1242
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Mahler, G., Richard, J., Clément, B., et al. 2018, *MNRAS*, **473**, 663
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *JChPh*, **21**, 1087
- Morishita, T., Abramson, L. E., Treu, T., et al. 2018, *ApJL*, **856**, L4
- Morishita, T., Abramson, L. E., Treu, T., et al. 2019, *ApJ*, **877**, 141
- Mowla, L., van Dokkum, P., Brammer, G., et al. 2018, arXiv:1808.04379
- Muzzin, A., Labbé, I., Franx, M., et al. 2012, *ApJ*, **761**, 142
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18
- Muzzin, A., Wilson, G., Lacy, M., Yee, H. K. C., & Stanford, S. A. 2008, *ApJ*, **686**, 966
- Neal, R. 2011, *MCMC Using Hamiltonian Dynamics* (Boca Raton, FL: CRC Press)
- Neal, R. M. 1993, *Probabilistic Inference Using Markov Chain Monte Carlo Methods* (Toronto: Univ. Toronto), <https://www.cs.toronto.edu/~radford/ftp/review.pdf>
- Nelson, D., Pillepich, A., Genel, S., et al. 2015, *A&C*, **13**, 12
- Newman, A., Treu, T., Ellis, R. S., et al. 2013a, *BAAS*, **45**, 203.02
- Newman, A. B., Belli, S., & Ellis, R. S. 2015, *ApJL*, **813**, L7
- Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018a, *ApJ*, **862**, 125
- Newman, A. B., Belli, S., Ellis, R. S., & Patel, S. G. 2018b, *ApJ*, **862**, 126
- Newman, A. B., Ellis, R. S., Andreon, S., et al. 2014, *ApJ*, **788**, 51
- Newman, A. B., Treu, T., Ellis, R. S., & Sand, D. J. 2013b, *ApJ*, **765**, 25
- Noll, S., Pierini, D., Cimatti, A., et al. 2009, *A&A*, **499**, 69
- Ocvirk, P., Pichon, C., Lançon, A., & Thiébaud, E. 2006, *MNRAS*, **365**, 46
- Panther, B., Jimenez, R., Heavens, A. F., & Charlot, S. 2007, *MNRAS*, **378**, 1550
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2011, *GALFIT: Detailed Structural Decomposition of Galaxy Images v3.0.5*, Astrophysics Source Code Library, ascl:1104.010
- Richard, J., Smith, G. P., Kneib, J.-P., et al. 2010, *MNRAS*, **404**, 325
- Rigby, J., Gladders, M., Sharon, K., & Wuyts, E. 2012, *sptz prop.*, 90232
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, *PeerJ*, **4**, e1686
- Sethuraman, J. 1994, *Statistica Sinica*, **4**, 639
- Sharon, K., Bayliss, M. B., Dahle, H., et al. 2020, *ApJS*, **247**, 12
- Sharon, K., Gladders, M. D., Rigby, J. R., et al. 2012, *ApJ*, **746**, 161
- Siana, B., Teplitz, H. I., Chary, R.-R., Colbert, J., & Frayer, D. T. 2008, *ApJ*, **689**, 59
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, **214**, 24
- Smail, I., Swinbank, A. M., Richard, J., et al. 2007, *ApJL*, **654**, L33
- Smith, R. J., Lucey, J. R., & Conroy, C. 2015, *MNRAS*, **449**, 3441
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, **214**, 15
- Stark, D. P., Swinbank, A. M., Ellis, R. S., et al. 2008, *Natur*, **455**, 775
- Strateva, I., Ivezić, Ž., Knapp, G. R., et al. 2001, *AJ*, **122**, 1861
- Swinbank, A. M., Webb, T. M., Richard, J., et al. 2009, *MNRAS*, **400**, 1121
- Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, *ApJ*, **749**, 121
- Tacchella, S., Carollo, C. M., Renzini, A., et al. 2015, *Sci*, **348**, 314
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, *MNRAS*, **458**, 242
- Toft, S., Zabl, J., Richard, J., et al. 2017, *Natur*, **546**, 510
- Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, *MNRAS*, **381**, 1252
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, **788**, 28
- van Dokkum, P. G., & Brammer, G. 2010, *ApJL*, **718**, L73
- van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, **677**, L5
- Verdugo, T., Motta, V., Muñoz, R. P., et al. 2011, *A&A*, **527**, A124
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, *MNRAS*, **444**, 1518
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014b, *Natur*, **509**, 177
- Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, *MNRAS*, **449**, 361
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014a, *ApJ*, **795**, 104
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, **735**, 86
- Whitaker, K. E., Rigby, J. R., Brammer, G. B., et al. 2014b, *ApJ*, **790**, 143
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, *ApJL*, **770**, L39
- Williams, C. C., Giavalisco, M., Bezanson, R., et al. 2017, *ApJ*, **838**, 94
- Williams, C. C., Giavalisco, M., Cassata, P., et al. 2014, *ApJ*, **780**, 1
- Williams, L. L. R., & Lewis, G. F. 1996, *MNRAS*, **281**, L35
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, **655**, 51
- Yee, H. K. C., Ellingson, E., Bechtold, J., Carlberg, R. G., & Cuillandre, J. C. 1996, *AJ*, **111**, 1783